

## FIGURE 1A

### 20CAGHA

CTGCAGGCCAGCGTCCTGATAAGTGAATTGCGCCGCCACCATGGGAGGCCCCAC  
CGTCAACCCCCCAGCAGCAGCAACAGCAGCAGCAACAGCAACAGCAGCAGC  
AACAACAGCAGCAGCAACAGACTAGTCGTACGTATCCCTATGACGTGCCCGA  
CTATGCGTAG

### 127CAGHA

CTGCAGGCCAGCGTCCTGATAAGTGAATTGCGCCGCCACCATGGGAGGCCCCAC  
CGTCAACCCCCCAGCAGCAGCAACAGCAGCAGCAACAGCAACAGCAGCAGC  
AACAACAGCAGCAGCAACAGCAACAGCAGCAGCAACAGCAACAGCAGCAGCAAC  
AGCAACAGCAGCAGCAACAGCAGCAGCAACAGCAACAGCAGCAGCAACAGCAAC  
AGCAGCAGCAACAGCAACAGCAGCAGCAACAGCAACAGCAACAGCAACAGCAAC  
AGCAGCAGCAACAGCAGCAGCAACAGCAACAGCAGCAGCAACAGCAACAGCAAC  
AGCAACAGCAACAGCAGCAGCAACAGCAACAGCAGCAGCAACAGCAACAGCAGC  
AGCAACAGCAGCAGCAACAGCAACAGCAGCAGCAACAGCAACAGCAGCTGCAAC  
AGCAACAGCAGCAGCAACAGCAACAGCAGCAGCAACAGCAACAGCAACAGCAAC  
CCTATGACGTGCCCGACTATGCGTAG

## FIGURE 1B

### 20QHA

MGGPPSTPQ<sub>20</sub>*TSRTYPYDV*PDYA

### 127QHA

MGGPPSTPQ<sub>127</sub>*TSRTYPYDV*PDYA

Figure 1. A) DNA sequences of 20QHA and 127QHA and B) their predicted protein sequences. The protein-coding region is underlined. The Kozak sequence is in italic.

FIGURE 2

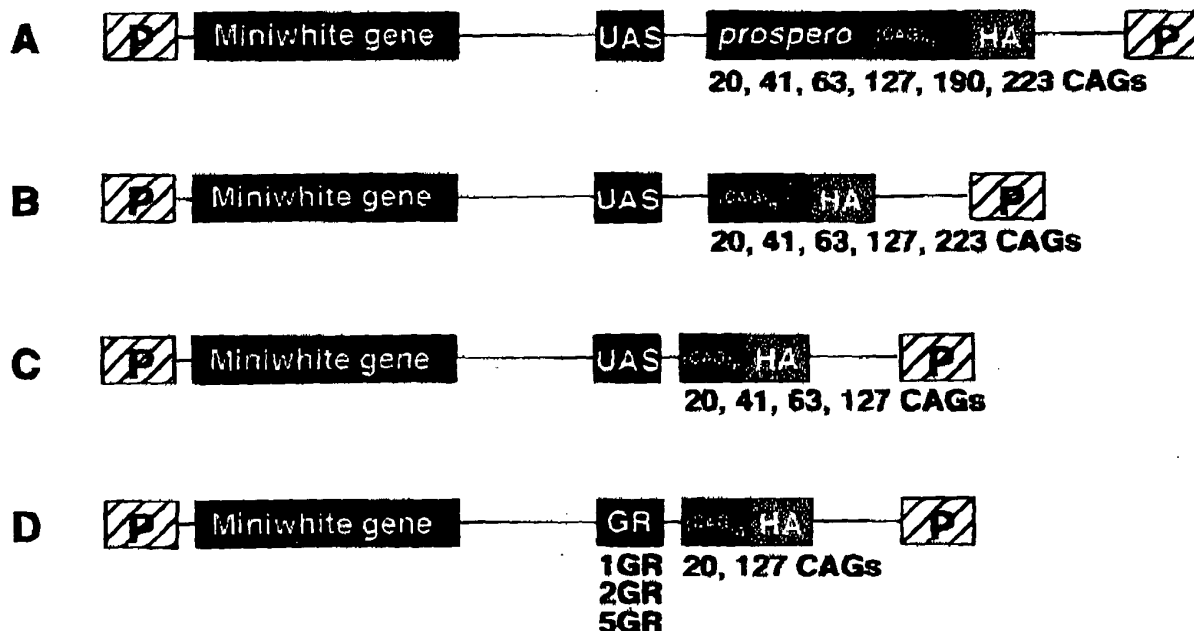


Figure 1. P-element plasmid constructs for production of transgenic flies. Each construct has two P-elements for chromosomal insertion. To facilitate identification of transformed flies, a miniwhite gene is included to produce red pigmentation in the eye. **A)** Plasmids carrying the full-length cDNA encoding the fly PROSPERO with various CAG repeat sizes. The expression of PROSPERO is regulated by five tandem upstream activating sequences (UAS). The yeast transcription factor GAL4 activates the transcription from these UAS elements. At its 3'-end, *prospero* cDNA is joined, in-frame, to a short DNA sequence that codes for a heterologous epitope, hemeagglutinin (HA). Antibodies against HA will be used to label the protein in immunohistochemical assays and Western blots. **B)** Plasmids carrying a partial cDNA encoding 422 amino acids of the C-terminal end of PROSPERO with various CAG repeat sizes. **C)** Plasmids carrying a DNA sequence that only encodes polyglutamines of various sizes. **D)** Plasmids carrying a DNA sequence that only encodes polyglutamines of various sizes, expressed under the control of one, two or five GLASS response elements (1GR, 2GR, or 5GR). The eye-specific protein GLASS activates the expression of polyglutamines from the GLASS response elements.

## FIGURE 3

### Generation of the P-element insertion and screening for modifiers

**M** P[Δ2-3]/P[Δ2-3] **X** **F** EP55/EP55

↓

**M** EP55/Y;; P[Δ2-3]/+ **X** **F** w/w

↓

**M** w/Y;pEP/+;+ **or** w/Y;+;pEP/+ **X** **F** w;GMR/CyO;127Q/127Q

↓

Progeny screened for eye phenotype

### Isolation of the new P-element insertion (pEP = suppressor or enhancer)

**M** (GMR;127Q)/pEP **X** **F** (CyO;TM3)/Xa

↓

**M** GMR/CyO;pEP/TM3 **X** **F** w1118

↓

**M** GMR;TM3 **or** CyO;pEP **X** **F** w;GMR/CyO;127Q/127Q to test

**X** **F** (CyO;TM3)/Xa to establish line

↓

**M** +/CyO;pEP/TM3 **X** **F** +/CyO;pEP/TM3

↓

pEP/TM3 **or** pEP/pEP established lines

Genetic scheme used for generating P-element mutants, screening for modifiers of polyglutamine toxicity, and isolating a hypothetical modifier P-element insertion on chromosome 3. Homozygous EP55 virgin females were crossed with males homozygous for a defective transposon, expressing the transposase. The F1 male progeny were crossed with virgin w1118 females. The F2 Male progeny that had coloured eyes and lacked the transposon's genetic markers were selected, as they contain a new stable insertion on an autosomal chromosome. These males were crossed with flies heterozygous for GMR-GAL4 on chromosome 2, balanced by CyO chromosome, and homozygous for UAS-127Q on chromosome 3. The resulting F3 progeny were screened for eye phenotype. Once a modifier was found, a single male was crossed to female (CyO;TM3)/Xa. The resulting male progeny were crossed to w1118 flies to separate the P-elements. This resulted in colored-eye progeny that carry a balancer for one chromosome and a P-element on another. Males from such progeny were tested for modifier activity by crossing to female w;GMR/CyO;127Q/127Q. The lines were established by crossing the latter males to (CyO;TM3)/Xa, and by crossing the resulting flies carrying CyO and TM3 balancers. **EP55**: source of transposable P-element; **P[Δ2-3]**: source of transposase; **F**: female; **M**: male; **CyO**: balancer chromosome 2; **TM3**: balancer chromosome 3. **Xa**: translocation (2;3) Xa. (Chromosome 4 is omitted.)

Genetic scheme used for generating P-element mutants, screening for modifiers of polyglutamine toxicity, and isolating a hypothetical modifier P-element insertion on chromosome 3.

FIGURE 4

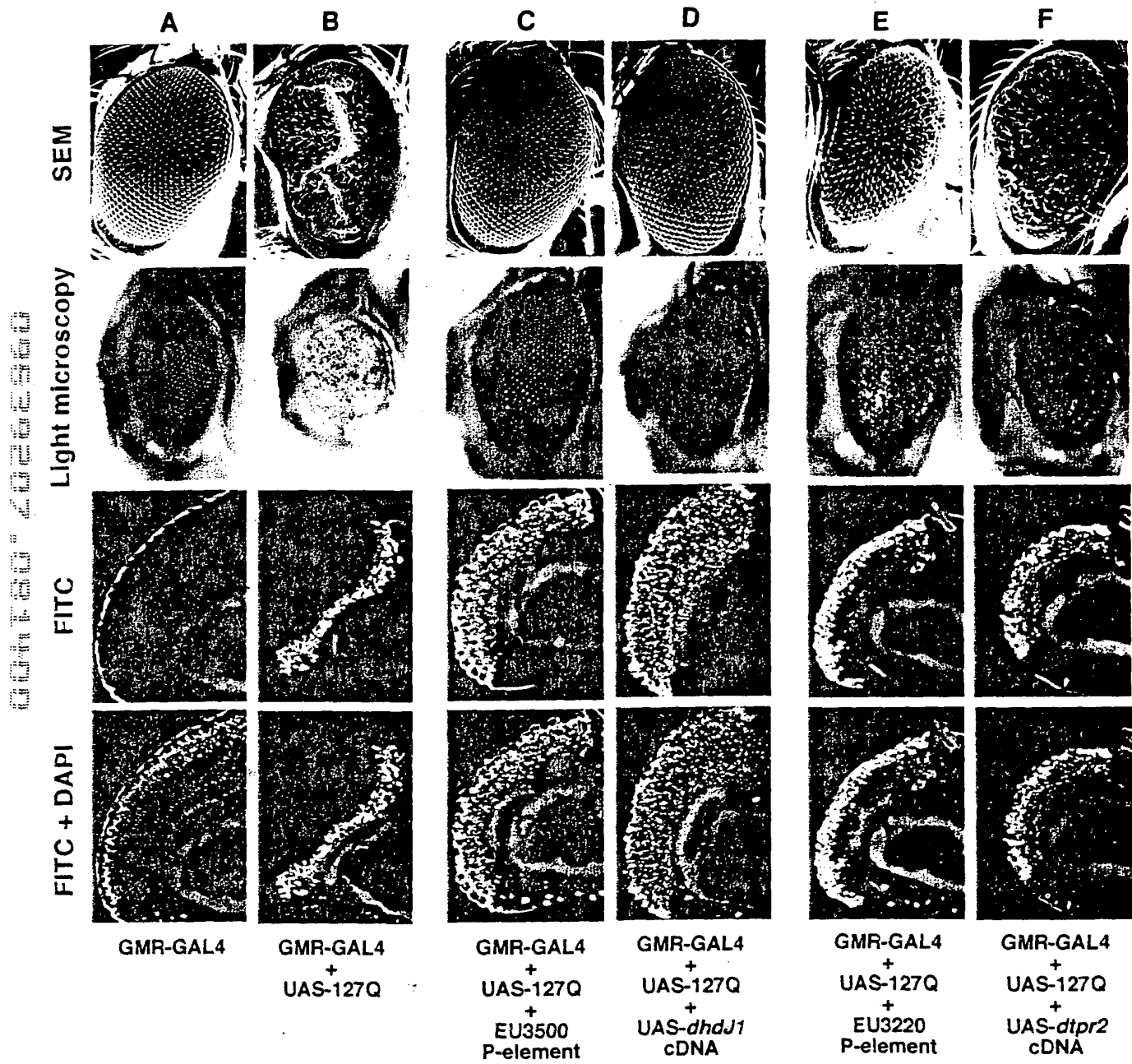
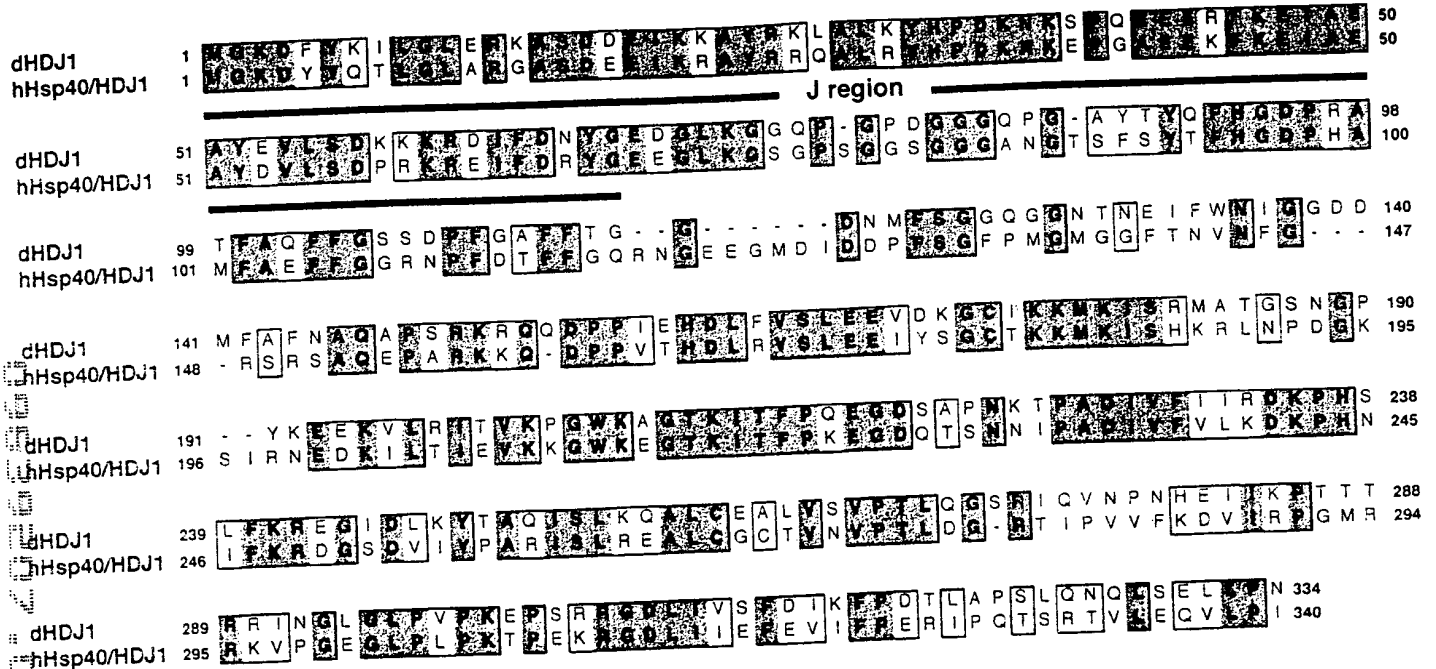
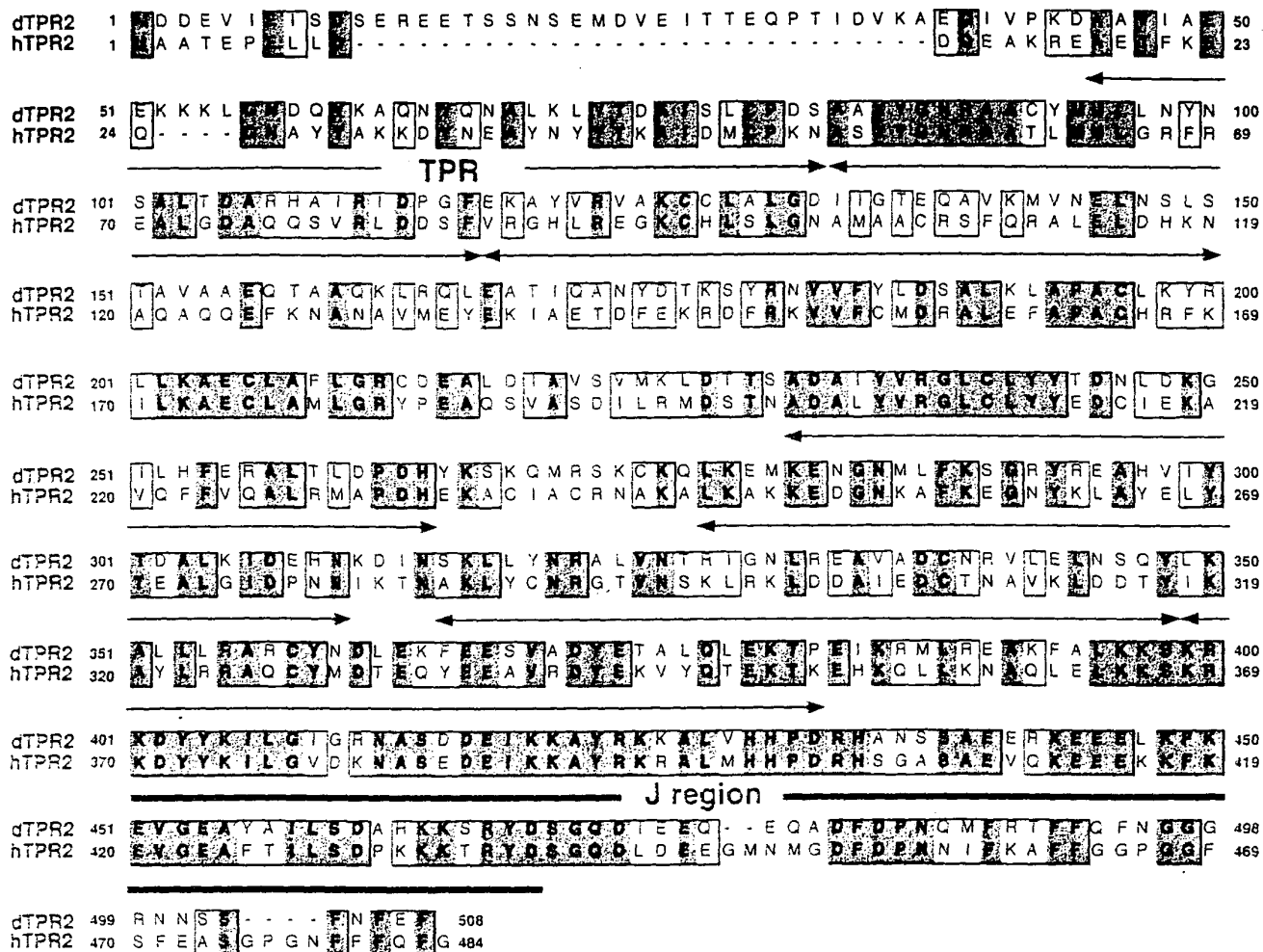




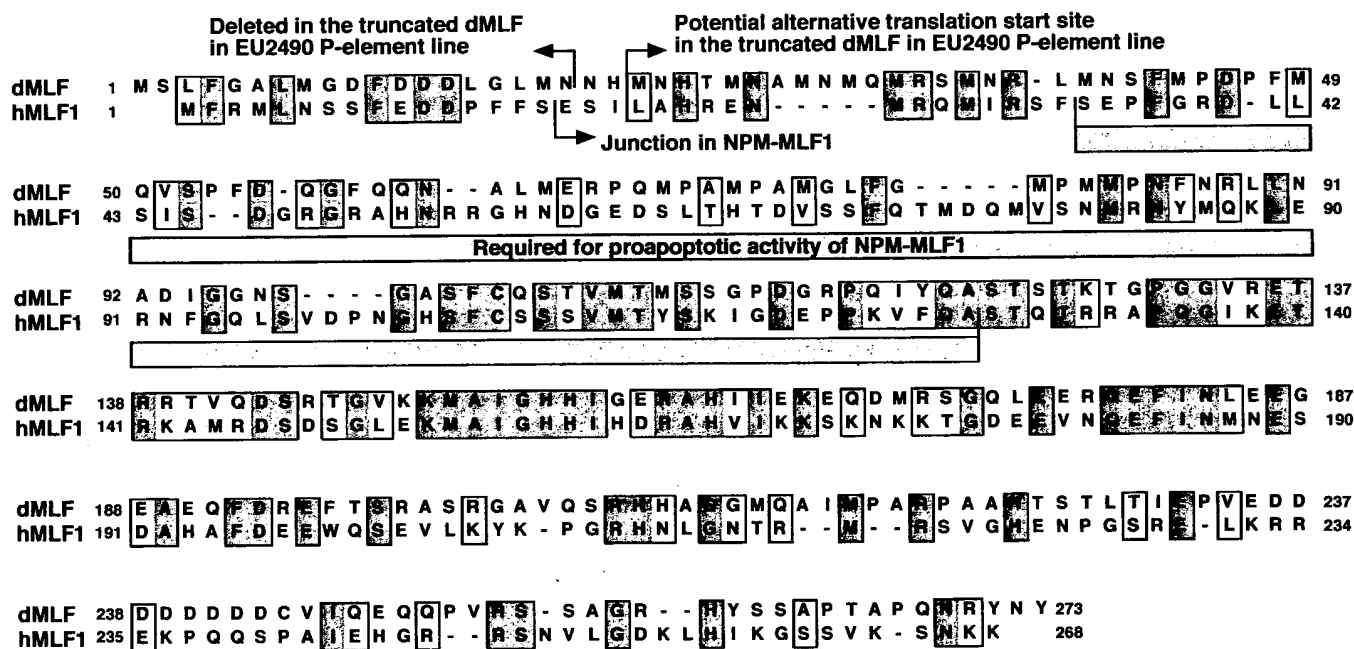
FIGURE 6



# FIGURE 7



# FIGURE 8



000000-000000

## FIGURE 9A

dTPR2 Protein 508 amino acids

MDDEVIEISDSEREETSSNSEMDVEITTEQPTIDVKAEQIVPKDAATIAEEKKKLG  
NDQYKAQNYQNALKLYTDAISLCPDSAAYYGNRAACYMMLLNYSALTDARH  
AIRIDPGFEKAYVRVAKCCALALGDIIGTEQAVKVMVNELNSLSTAVAAEQTAAQK  
LRQLEATIQANYDTKSYRNVVFYLD SALKLAPACLK YRLLKAECLAFLGRCDEA  
LDIAVSVMKLDTTSADAIYVRGLCLYYTDNLDKGILHFERALTLDPDHYKSKQM  
RSKCKQLKEMKENG NMLFKSGRYREAHVIYTDALKIDEHNKDINSKLLYNRALV  
NTRIGNLREAVADCNRVLELNSQYLKALLLRARCYNDLEKFEESVADYETALQL  
EKTPEIKRMLREAKFALKKSKRKDYKILGIGRNASDDEIKKAYRKALVHHPD  
RHANSSAEERKEEELKFKEVGEAYAILSDAHKKSR YDSGQDIEEQEQADFDPNQ  
MFR TFFQFN GGGGRNNSSFNFEF

## FIGURE 9B

dTPR2 cDNA 2239 base pairs

GGCACGAGCCACTACTTCGCATGGCACGCTTTTTTCCGTGTGCTCGGTTCGTT  
CGGCCATACAAAACACAAAATTCAAGTTTAAAACTAAATAGGCAACTAAAA  
GGGAAGCCGCAGCGAATAAAGTGATTTGCTGAAAGAGACGTAAGAAAAGTTA  
ATCGCATCGAAGGCACCAGAAATCGGGGATTTCTAACACGGCGCGCGTGCGA  
CGTACATACATACGCAAGCGCACACACACACGAACAATTACTTGCCATTGAC  
GCAAAAGCGAAAAAGCAGTGAATAAAGGGGAATTGACAAATAACAACGTT  
TTGCAAGCACTGGACTCTGGTCGCTGGTGTTCTTTCATTTTGTAATTGCCACG  
CATGGACGACGAAGTAATTGAAATTAGCGACAGCGAACGCGAAGAAACCTC  
ATCGAACTCCGAAATGGATGTGGAAATAACGACAGAACAGCCAACCATCGAT  
GTCAAAGCAGAGCAAATTGTGCCCAAGGACGCGGCAACCAATTGCCGAGGAG  
AAGAAGAACTGGGCAACGACCAATACAAGGCGCAGAACTATCAGAATGCA  
CTCAAGCTCTACACGGATGCCATATCGCTGTGTCCGGACTCGGCGGCATACTA  
TGGCAATCGGGCCGCTGCTACATGATGCTGCTCAACTATAATAGCGCCCTG  
ACCGACGCCCCGACACGCCATACGCATCGATCCGGGCTTCGAGAAGGCCTACG  
TCCGTGTGGCCAAGTGCTGTCTGGCCCTGGGCGACATTATTGGCACCGAACA  
GGCCGTCAAAAATGGTCAACGAGCTGAATTCGCTTAGCACGGCTGTTGCTGCC  
GAACAGACGGCGGCGCAAAAAGTTGCGCCAATTGGAGGCCACCAATTCAGGCG  
AACTACGATACGAAATCCTATCGCAATGTGGTCTTCTATTTGGATAGTGCCTT  
GAAATTGGCGCCCCGCTGTTTGAAATATCGTCTACTCAAGGCTGAGTGCCTTG  
CATTTTTGGGGCGATGTGATGAGGCCTTGACATTGCGGTCAGTGTAATGAA  
ACTGGATACCACATCGGCGGATGCGATATACGTGAGAGGTCTGTGCCTGTAC  
TACACGGACAACCTGGACAAGGGAATTCTTCATTTTCGAGCGCGCCCTGACCC  
TCGACCCGGACCACTACAAGTCCAAGCAGATGCGCAGCAAATGCAAGCAGCT  
CAAGGAGATGAAGGAGAACGGCAATATGCTATTCAAGTCGGGTTCGGTATCGC  
GAGGCACACGTTATCTACACGGACGCCCTGAAGATCGATGAACACAACAAGG  
ATATCAATTCGAAATTGCTTTACAATCGGGCTTTGGTCAACACGCGTATTGGC  
AATTTGCGAGAGGCCGTGGCCGATTGCAATCGAGTGCTGGAGCTGAATAGTC  
AGTATCTGAAGGCTCTGTTGCTGCGAGCGCGCTGCTACAATGATCTGGAGAA  
GTTTCGAGGAGTCGGTGGCGGACTATGAGACGGCGCTGCAGCTGGAGAAGAC  
GCCGGAGATTAAGCGAATGCTGCGCGAGGCCAAGTTTGCGTTGAAGAAGTCG  
AAGCGAAAGGACTACTACAAGATCCTGGGCATTGGACGCAATGCGTCCGACG  
ACGAGATCAAGAAGGCGTATCGCAAAAAGGCGCTGGTACATCATCCGGATCG

## FIGURE 9B (Continued)

ACACGCAAACAGCAGTGCCGAGGAGCGCAAGGAGGAGGAGCTCAAGTTCAA  
GGAGGTGGGCGAGGCGTACGCCATACTGTCCGATGCTCACAAGAAGTCGCGC  
TACGACAGCGGCCAGGATATCGAGGAGCAGGAGCAAGCCGACTTCGATCCG  
AATCAAATGTTCCGCACATTCTTCCAATTCAACGGCGGTGGCCGGAATAATC  
ATCGTTCAACTTTGAGTTCTAGGATCCCAACGAGTGTTGTTCAACCACCACAGA  
GAAGAAGACCATCTCAATCCCATACTTTCTGCCTCATCCGAAACCAACATAC  
AGCAGCGCACAAATTTTGAACCTCTTTTACATATTTCTTTTCCAAAAAGCAAGA  
AAATACCACATTTTGATTATGTTAACGAATGAATATATGCCAAGTTATTTGAA  
AAAATATTCTAAATCAAATAATGCAACTAAATTTCCAGTGTAAGTTCACATT  
TTTAAATGTTCTTTCTTGGATTTTTTTTTTCGGCAACATTAATAAATCATGGGAG  
ATTTGTGTAAATAAACAGAAATATACATATAAAAAAAAAAAAAAAAAAAAAA

001100-20000000

## FIGURE 10A

dMLF Protein 273 amino acids

MSLFGALMGDFDDDLGLMNNHNMNHTMNAMNMQMRSMNRLMNSFMPDPFMQ  
VSPFDQGFQQNALMERPQMPAMPAMGLFGMPMPNPNRLLNADIGGNSGASF  
CQSTVMTMSSGPDGRPQIYQASTSTKTGPGGVRETRRTVQDSRTGVKKMAIGHH  
IGERAHIIIEKEQDMRSGQLEERQEFINLEEAEQFDREFTSRASRGAVQSRHHAG  
GMQAIMPARPAHTSTLTIEPVEDDDDDDDDCVIEQQPVRSSAGRHYSSAPTAP  
QNRNYN

## FIGURE 10B

dMLF cDNA 1753 base pairs

GGCACGAGGAAAATATTCGTGAAAATTCTGCATACGGAAAGAAGAAAATTC  
GAGCAACAGAAAGCCAACACAATCCACAAAAATGTCTTTATTCGGAGCGTTG  
ATGGGTGATTTTCGACGACGATCTCGGCCTTATGAACAACCACATGAACCACA  
CTATGAACGCGATGAACATGCAGATGCGCTCGATGAATCGCCTGATGAACAG  
CTTTATGCCCCGATCCCTTCATGCAGGTCTCGCCCTTTGACCAGGGATTCCAGC  
AGAACGCTCTCATGGAGCGTCCGCAGATGCCGGCCATGCCAGCCATGGGCCT  
CTTCGGCATGCCCATGATGCCAACTTTAATCGCCTGTTGAACGCTGATATTG  
GTGGCAATTCAGGCGCATCCTTCTGCCAGAGCACCGTGATGACCATGTCATC  
GGGTCCCGATGGGCGTCCTCAGATCTACCAGGCCAGCACTAGTACCAAAACA  
GGACCGGGAGGCGTTCTGTGAGACCCGCAGGACGGTGCAGGACTCGCGCACT  
GGGGTGAAGAAGATGGCCATTGGTCATCACATCGGCGAGCGGGCACACATTA  
TTGAGAAAGAGCAGGACATGCGCTCAGGACAACCTGGAGGAGCGCCAGGAGT  
TCATTAATCTGGAGGAGGGAGAAGCCGAGCAGTTTGACAGGGAGTTTACATC  
GCGCGCTAGTCGCGGAGCGGTGCAGTCAAGACATCATGCTGGTGGCATGCAG  
GCCATCATGCCCCGCCGTCCAGCGGCACACACCTCGACGTTGACCATTGAGC  
CAGTGGAGGACGACGACGACGATGATGATGACTGTGTAATCCAGGAGCAGC  
AACCGGTTCTGCTCCTCCGCGGGCCGCCATTATTCCAGTGCGCCAACGGCACC  
GCAGAACAGATATAATTACTAAATCTAAAGTCAATACAGTATATTTTACTAA  
CTATCCGATAAAACAGAAACAGAATTGCATACTATAAATTTCTGCTAATTAC  
ATTCCCAACTGCGTTCAAACGAAACGAATATCGAATCGAAATCATAGAATGC  
ACAGAGCAGCATACATCCACATCCCTATGCCGCCAATCCGAGGCGCCAACAA  
CGTGCCGTAAAACATTTTACACGAGGACGAAGCGGCCAGCTCCTACAAGG  
CGGTCAAGCGCGGCAAGAAGAAGTAGTAGAAACGTGATCATCTGTATGCCAA  
CATCTTCCGCATCGCACACTCAAAAACACTAGGAAGCAAAGCGTTGGGTTCT  
GTTCCATAGCAGGAAAACCAATTCAAATATTTTTTAACAAACACAATTCTTTA  
CCAGTTCTGTCTTATCCTGCGTGAGTCGACCAGAATGCAACACTAAAAAATGT  
ACAACTTCAAGATGCTATTGATGTGCACGCAGGATACAGAACAACCTTGCTTA  
AATTTACTTAAAACAAATGTGACTATTCAACGCCGAAATCATTACAACACAC  
ACTCTCAGACCTAATCGAAAAATTCAATGAAAGTAATGGAATATATATGAAA  
TCGTAATTATAAGTTTGAATTATTTGATTAATTCTCAAGTTTTTAGATTTTGT  
AGCCACTAAGCTTTAAATTATGGATGCCAGTTAGCGTGCAAATGAACACAAT  
TGATTTGAAGGCTCCGAACGATAGAAAACAACAATTACCAATTCCCCAAATA  
CATGTAATTCGTAAGGCCTAAGTAAATGTTAACGTGAATTTAATTAAATGGTA  
ATTACATTATAATAGTAAAAAAAAAAAAAAAAAAAA

### ClustalW Formatted Alignments





[illegible]

AA950779  
A1458556  
A1457049  
A1110008  
AA201154  
AA952015  
A1516974  
AA201429  
AA392620  
A1439655  
A1544064  
A1516760  
AA978757  
AA202799  
AA735495  
AA201527  
AA438619  
A1516274  
AA440018  
AA820977  
A1555655

dTPR2 (GH09432) GenBank submit

420 G A A A T A A C G A C A G A A C A G C C A A C C A T C G A T G T C A A A G C A G A G C A A A T T G T 440  
274  
424 G A A A T A A C G A C A G A A C A G C C A A C C A T C G A T G T C A A A G C A G A G C A A A T T G T 460  
428 G A A A T A A C G A C A G A A C A G C C A A C C A T C G A T G T C A A A G C A G A G C A A A T T G T 480  
432 G A A A T A A C G A C A G A A C A G C C A A C C A T C G A T G T C A A A G C A G A G C A A A T T G T 500  
436 G A A A T A A C G A C A G A A C A G C C A A C C A T C G A T G T C A A A G C A G A G C A A A T T G T 520  
387  
412 A T A T C C T G C C C G C T G T C A T A G C G C G A G T T C T T G T G A G C A T C C G A C A G T A T 560  
407 G A A A T A A C G A C A G A A C A G C C A A C C A T C G A T G T C A A A G C A G A G C A A A T T G T 580  
403 G A A A T A A C G A C A G A A C A G C C A A C C A T C G A T G T C A A A G C A G A G C A A A T T G T 600  
410 G A A A T A A C G A C A G A A C A G C C A A C C A T C G A T G T C A A A G C A G A G C A A A T T G T 620  
406 G A A A T A A C G A C A G A A C A G C C A A C C A T C G A T G T C A A A G C A G A G C A A A T T G T 640  
388  
402 G A A A T A A C G A C A G A A C A G C C A A C C A T C G A T G T C A A A G C A G A G C A A A T T G T 660  
378  
418 G A A A T A A C G A C A G A A C A G C C A A C C A T C G A T G T C A A A G C A G A G C A A A T T G T 700  
414 G A A A T A A C G A C A G A A C A G C C A A C C A T C G A T G T C A A A G C A G A G C A A A T T G T 720  
410 G A A A T A A C G A C A G A A C A G C C A A C C A T C G A T G T C A A A G C A G A G C A A A T T G T 740  
406 G A A A T A A C G A C A G A A C A G C C A A C C A T C G A T G T C A A A G C A G A G C A A A T T G T 760  
374  
418 G A A A T A A C G A C A G A A C A G C C A A C C A T C G A T G T C A A A G C A G A G C A A A T T G T 800  
414 G A A A T A A C G A C A G A A C A G C C A A C C A T C G A T G T C A A A G C A G A G C A A A T T G T 820  
410 G A A A T A A C G A C A G A A C A G C C A A C C A T C G A T G T C A A A G C A G A G C A A A T T G T 840  
406 G A A A T A A C G A C A G A A C A G C C A A C C A T C G A T G T C A A A G C A G A G C A A A T T G T 860  
372

AA950779  
A1456556  
A1457049  
A1110008  
AA201154  
AA952015  
A1516974  
AA2011429  
AA392620  
AA439655  
A1544064  
A1516760  
AA978757  
AA202799  
AA735495  
AA201527  
AA438619  
A1516274  
AA440018  
AA820977  
AA525605  
dTPR2 (GH9432) GenBank submit

AA950779 520 ACCAATACAAAGGCGCAGAACTATC AGAATGCACCTCAAGCTCTACACGGAT 569  
AI456556 524 ACCAATACAAAGGCGCAGAACTATC AGAATGCACCTCAAGCTCTACACGGAT 573  
AI457049 524 ACCAATACAAAGGCGCAGAACTATC AGAATGCACCTCAAGCTCTACACGGAT 566  
AI110008 527 ACCAATACAAAGGCGCAGAACTATC AGAATGCACCTCAAGCTCTACACGGAT 576  
AA201154 529 ACCAATACAAAGGCGCAGAACTATC AGAATGCACCTCAAGCTCTACACGGAT 570  
AA952015 527 ACCAATACAAAGGCGCAGAACTATC AGAATGCACCTCAAGCTCTACACGGAT 576  
AI516974 535 ACCAATACAAAGGCGCAGAACTATC AGAATGCACCTCAAGCTCTACACGGAT 584  
AA201429 537 ACCAATACAAAGGCGCAGAACTATC AGAATGCACCTCAAGCTCTACACGGAT 586  
AA392620 512 CCTCGGGCACTCTCTTTTGGGCTG AGAATGCACCTCAAGCTCTACACGGAT 534  
AA439655 508 ACCAATACAAAGGCGCAGAACTATC AGAATGCACCTCAAGCTCTACACGGAT 555  
AI544064 516 ACCAATACAAAGGCGCAGAACTATC AGAATGCACCTCAAGCTCTACACGGAT 559  
AI516760 510 ACCAATACAAAGGCGCAGAACTATC AGAATGCACCTCAAGCTCTACACGGAT 531  
AA978757 490 ACCAATACAAAGGCGCAGAACTATC AGAATGCACCTCAAGCTCTACACGGAT 479  
AA202799 368 ACCAATACAAAGGCGCAGAACTATC AGAATGCACCTCAAGCTCTACACGGAT 367  
AA735495 504 ACCAATACAAAGGCGCAGAACTATC AGAATGCACCTCAAGCTCTACACGGAT 503  
AA201527 371 ACCAATACAAAGGCGCAGAACTATC AGAATGCACCTCAAGCTCTACACGGAT 370  
AA438619 378 ACCAATACAAAGGCGCAGAACTATC AGAATGCACCTCAAGCTCTACACGGAT 377  
AI516274 519 ACCAATACAAAGGCGCAGAACTATC AGAATGCACCTCAAGCTCTACACGGAT 568  
AA440018 513 ACCAATACAAAGGCGCAGAACTATC AGAATGCACCTCAAGCTCTACACGGAT 562  
AA820977 519 ACCAATACAAAGGCGCAGAACTATC AGAATGCACCTCAAGCTCTACACGGAT 564  
AI455565 534 CCTCGGGCACTCTCTTTTGGGCTG AGAATGCACCTCAAGCTCTACACGGAT 555  
dTPR2 (GH09432) GenBank submit 508 ACCAATACAAAGGCGCAGAACTATC AGAATGCACCTCAAGCTCTACACGGAT 587

# FIGURE 11A (Continued)

AA950779 570 GCCATATCGCT 540  
 A1456556 274 573  
 A1457049 586 545  
 A1110008 577 GCCATATCGCTGTGTCCGGACTCGGCGGCATACTATGGCAATCGGGCCG 525  
 AA201154 530 520  
 AA952015 577 GCCATATCGCTGTGTCCGGACTCGGCGGCATACTATGGCAATCGGGCCGC 525  
 A1516974 586 GCCATATCGCTGTGTCCGGACTCGGCGGCATACTATGGCAATCGGGCCGC 534  
 AA201429 387 546  
 AA392620 535 534  
 AA439655 556 GCCATATCGCTGTG 549  
 A1544064 580 GCCATATCGCTGTGTCCGGACTCGGCGGCATACTATGGCAATCGGGCCGC 529  
 A1516760 532 531  
 AA978757 480 479  
 AA202799 368 367  
 AA735495 504 523  
 AA201527 371 370  
 AA438619 378 377  
 A1516274 589 GCCATATCGCTGTGTCCGGACTCGGCGGCATACTATGGCAATCGGGCCGC 619  
 AA440018 583 GCCATATCGCTGTGTCCGGACTCGGCGGCATACTATGGCAATCGGGCCGC 613  
 AA820977 589 GCCATATCGCTGTGTCCGGA 548  
 A1455565 586 545  
 dTPR2 (GH09432) GenBank submit 589 GCCATATCGCTGTGTCCGGACTCGGCGGCATACTATGGCAATCGGGCCGC 637

AA950779 581 540  
 A1456556 274 573  
 A1457049 586 545  
 A1110008 626 CTGCTACATGATGCTGCTCAACTATAATAGCGCCCTGACCGACGCCCGAC 675  
 AA201154 530 520  
 AA952015 627 CTGCTACATGATGCTGCTCAACTATAATAGCGCCCTGACCGACGCCCGAC 676  
 A1516974 626 CTGCTACATGATGCTGCTCAACTATAATAGCGCCCTGACCGAC 677  
 AA201429 387 546  
 AA392620 535 534  
 AA439655 570 549  
 A1544064 610 CTGCTACATGATGCTGCTCAACTATAATAGCGCCCTGACCGACGCCCGAC 651  
 A1516760 532 531  
 AA978757 480 479  
 AA202799 368 367  
 AA735495 504 523  
 AA201527 371 370  
 AA438619 378 377  
 A1516274 619 CTGCTACATGATGCTGCTCAACTATAATAGCGCCCTGACCGACGCCCGAC 649  
 AA440018 613 CTGCTACATGATGCTGCTCAACT 615  
 AA820977 589 548  
 A1455565 586 545  
 dTPR2 (GH09432) GenBank submit 638 CTGCTACATGATGCTGCTCAACTATAATAGCGCCCTGACCGACGCCCGAC 687

AA950779 581 540  
 A1456556 274 573  
 A1457049 586 545  
 A1110008 676 ACG 679  
 AA201154 530 520  
 AA952015 677 ACGCCATACGCATCGATCCGGGCTTCGAGAAGGCCCTACGTCCGTGTGGCC 726  
 A1516974 678 677  
 AA201429 387 546  
 AA392620 535 534  
 AA439655 570 549  
 A1544064 680 ACGCCATACGCATCGATCCGGGCTTCGAG 684  
 A1516760 532 531  
 AA978757 480 479  
 AA202799 368 367  
 AA735495 504 523  
 AA201527 371 370  
 AA438619 378 377  
 A1516274 689 ACGCCATACGCATCGATCCGGGC 691  
 AA440018 636 635  
 AA820977 589 548  
 A1455565 586 545  
 dTPR2 (GH09432) GenBank submit 688 ACGCCATACGCATCGATCCGGGCTTCGAGAAGGCCCTACGTCCGTGTGGCC 737

AA950779 581 540  
 A1456556 274 573  
 A1457049 586 545  
 A1110008 679 678  
 AA201154 530 520  
 AA952015 727 AAGTGCTGTCTGGCCCTGGGCGACATTATTGGCCCGAACAGGCCGTCAA 776  
 A1516974 678 677  
 AA201429 387 546  
 AA392620 535 534  
 AA439655 570 549  
 A1544064 689 688  
 A1516760 532 531  
 AA978757 480 479  
 AA202799 368 367  
 AA735495 504 523  
 AA201527 371 370  
 AA438619 378 377  
 A1516274 692 691  
 AA440018 636 635  
 AA820977 589 548  
 A1455565 586 545  
 dTPR2 (GH09432) GenBank submit 738 AAGTGCTGTCTGGCCCTGGGCGACATTATTGGCACCGAACAGGCCGTCAA 787

# FIGURE 11A (Continued)

AA950779	581		580
AI456556	274		273
AI457049	586		565
AI110008	679		678
AA201154	530		529
AA952015	777	A T G G T T	782
AI516974	678		677
AA201429	367		386
AA392620	535		534
AA439655	570		569
AI544064	689		688
AI516760	532		531
AA978757	480		479
AA202799	368		367
AA735495	504		503
AA201527	371		370
AA438619	378		377
AI516274	692		691
AA440018	636		635
AA820977	589		588
AI455565	586		565
dTPR2 (GH09432) GenBank submit	788	A A T G G T C A A C G A G C T G A A T T C G C T T A G C A C G G C T G T T G C T G C C G A A C A G A	837

AA950779	581		580
AI456556	274		273
AI457049	586		565
AI110008	679		678
AA201154	530		529
AA952015	783		782
AI516974	678		677
AA201429	367		386
AA392620	535		534
AA439655	570		569
AI544064	689		688
AI516760	532		531
AA978757	480		479
AA202799	368		367
AA735495	504		503
AA201527	371		370
AA438619	378		377
AI516274	692		691
AA440018	636		635
AA820977	589		588
AI455565	586		565
dTPR2 (GH09432) GenBank submit	838	C G G C G G C G C A A A A G T T G C G C C A A T T G G A G G C C A C C A T T C A G G C G A A C T A C	887

AA950779	581		580
AI456556	274		273
AI457049	586		565
AI110008	679		678
AA201154	530		529
AA952015	783		782
AI516974	678		677
AA201429	367		386
AA392620	535		534
AA439655	570		569
AI544064	689		688
AI516760	532		531
AA978757	480		479
AA202799	368		367
AA735495	504		503
AA201527	371		370
AA438619	378		377
AI516274	692		691
AA440018	636		635
AA820977	589		588
AI455565	586		565
dTPR2 (GH09432) GenBank submit	888	G A T A C G A A A T C C T A T C G C A A T G T G G T C T T C T A T T T G G A T A G T G C C T T G A A	937

AA950779	581		580
AI456556	274		273
AI457049	586		565
AI110008	679		678
AA201154	530		529
AA952015	783		782
AI516974	678		677
AA201429	367		386
AA392620	535		534
AA439655	570		569
AI544064	689		688
AI516760	532		531
AA978757	480		479
AA202799	368		367
AA735495	504		503
AA201527	371		370
AA438619	378		377
AI516274	692		691
AA440018	636		635
AA820977	589		588
AI455565	586		565
dTPR2 (GH09432) GenBank submit	938	A T T G G C G C C C G C C T G T T T G A A A T A T C G T C T A C T C A A G G C T G A G T G C C T T G	987

GenBank accession numbers

# FIGURE 11A (Continued)

AA950779	581	580
AI456556	274	273
AI457049	566	565
AI110008	679	678
AA201154	530	529
AA952015	783	782
AI516974	678	677
AA201429	367	366
AA392620	535	534
AA439655	570	569
AI544064	689	688
AI516760	532	531
AA978757	480	479
AA202799	368	367
AA735495	504	503
AA201527	371	370
AA438619	378	377
AI516274	682	691
AA440018	635	635
AA820977	589	588
AI455565	566	565
dTPR2 (GH09432) GenBank submit	588	1037

C A T T T T T T G G G G C G A T G T G A T G A G G C C T T G G A C A T T G C G G T C A G T G T A A T G

AA950779	581	580
AI456556	274	273
AI457049	566	565
AI110008	679	678
AA201154	530	529
AA952015	783	782
AI516974	678	677
AA201429	367	366
AA392620	535	534
AA439655	570	569
AI544064	689	688
AI516760	532	531
AA978757	480	479
AA202799	368	367
AA735495	504	503
AA201527	371	370
AA438619	378	377
AI516274	682	691
AA440018	635	635
AA820977	589	588
AI455565	566	565
dTPR2 (GH09432) GenBank submit	1038	1087

A A A C T G G A T A C C A C A T C G G C G G A T G C G A T A T A C G T G A G A G G T C T G T G C C T

AA950779	581	580
AI456556	274	273
AI457049	566	565
AI110008	679	678
AA201154	530	529
AA952015	783	782
AI516974	678	677
AA201429	367	366
AA392620	535	534
AA439655	570	569
AI544064	689	688
AI516760	532	531
AA978757	480	479
AA202799	368	367
AA735495	504	503
AA201527	371	370
AA438619	378	377
AI516274	682	691
AA440018	635	635
AA820977	589	588
AI455565	566	565
dTPR2 (GH09432) GenBank submit	1088	1137

G T A C T A C A C G G A C A A C C T G G A C A A G G G A A T T C T T C A T T T C G A G C G C G C C C

AA950779	581	580
AI456556	274	273
AI457049	566	565
AI110008	679	678
AA201154	530	529
AA952015	783	782
AI516974	678	677
AA201429	367	366
AA392620	535	534
AA439655	570	569
AI544064	689	688
AI516760	532	531
AA978757	480	479
AA202799	368	367
AA735495	504	503
AA201527	371	370
AA438619	378	377
AI516274	682	691
AA440018	635	635
AA820977	589	588
AI455565	566	565
dTPR2 (GH09432) GenBank submit	1138	1187

T G A C C C T C G A C C C G G A C C A C T A C A A G T C C A A G C A G A T G C G C A G C A A A T G C

# FIGURE 11A (Continued)

AA950779	581	580
AI456556	274	273
AI457049	585	565
AI110008	679	678
AA201154	530	529
AA952015	783	782
AI516974	678	677
AA201429	367	386
AA392620	535	534
AA439655	570	569
AI544064	689	688
AI516760	532	531
AA978757	480	479
AA202799	368	367
AA735495	504	503
AA201527	371	370
AA438619	378	377
AI516274	692	691
AA440018	636	635
AA820977	589	588
AI455565	585	565
dTPR2 (GH09432) GenBank submit	1188	1237

A A G C A G C T C A A G G A G A T G A A G G A G A A C G G C A A T A T G C T A T T C A A G T C G G G

AA950779	581	580
AI456556	274	273
AI457049	585	565
AI110008	679	678
AA201154	530	529
AA952015	783	782
AI516974	678	677
AA201429	367	386
AA392620	535	534
AA439655	570	569
AI544064	689	688
AI516760	532	531
AA978757	480	479
AA202799	368	367
AA735495	504	503
AA201527	371	370
AA438619	378	377
AI516274	692	691
AA440018	636	635
AA820977	589	588
AI455565	585	565
dTPR2 (GH09432) GenBank submit	1238	1287

T C G G T A T C G C G A G G C A C A C G T T A T C T A C A C G G A C G C C C T G A A G A T C G A T G

AA950779	581	580
AI456556	274	273
AI457049	585	565
AI110008	679	678
AA201154	530	529
AA952015	783	782
AI516974	678	677
AA201429	367	386
AA392620	535	534
AA439655	570	569
AI544064	689	688
AI516760	532	531
AA978757	480	479
AA202799	368	367
AA735495	504	503
AA201527	371	370
AA438619	378	377
AI516274	692	691
AA440018	636	635
AA820977	589	588
AI455565	585	565
dTPR2 (GH09432) GenBank submit	1288	1337

A A C A C A A C A A G G A T A T C A A T T C G A A A T T G C T T T A C A A T C G G G C T T T G G T C

AA950779	581	580
AI456556	274	273
AI457049	585	565
AI110008	679	678
AA201154	530	529
AA952015	783	782
AI516974	678	677
AA201429	367	386
AA392620	535	534
AA439655	570	569
AI544064	689	688
AI516760	532	531
AA978757	480	479
AA202799	368	367
AA735495	504	503
AA201527	371	370
AA438619	378	377
AI516274	692	691
AA440018	636	635
AA820977	589	588
AI455565	585	565
dTPR2 (GH09432) GenBank submit	1338	1387

A A C A C G C G T A T T G G C A A T T T G C G A G A G G C C G T G G C C G A T T G C A A T C G A G T

# FIGURE 11A (Continued)

AA950779	581	580
A1456556	274	273
A1457049	586	565
A1110008	679	678
AA201154	530	529
AA952015	783	782
A1516974	678	677
AA201429	367	386
AA392620	535	534
AA439655	570	569
A1544064	689	688
A1516760	532	531
AA978757	480	479
AA202799	368	367
AA735495	504	503
AA201527	371	370
AA438619	378	377
A1516274	682	691
AA440018	636	635
AA820977	589	588
A1455565	586	565
dTPR2 (GH09432) GenBank submit	1388	1437

G C T G G A G C T G A A T A G T C A G T A T C T G A A G G C T C T G T T G C T G C G A G C G C G C T

AA950779	581	580
A1456556	274	273
A1457049	586	565
A1110008	679	678
AA201154	530	529
AA952015	783	782
A1516974	678	677
AA201429	367	386
AA392620	535	534
AA439655	570	569
A1544064	689	688
A1516760	532	531
AA978757	480	479
AA202799	368	367
AA735495	504	503
AA201527	371	370
AA438619	378	377
A1516274	682	691
AA440018	636	635
AA820977	589	588
A1455565	586	565
dTPR2 (GH09432) GenBank submit	1438	1487

G C T A C A A T G A T C T G G A G A A G T T C G A G G A G T C G G T G G C G G A C T A T G A G A C G

AA950779	581	580
A1456556	274	273
A1457049	586	565
A1110008	679	678
AA201154	530	529
AA952015	783	782
A1516974	678	677
AA201429	367	386
AA392620	535	534
AA439655	570	569
A1544064	689	688
A1516760	532	531
AA978757	480	479
AA202799	368	367
AA735495	504	503
AA201527	371	370
AA438619	378	377
A1516274	682	691
AA440018	636	635
AA820977	589	588
A1455565	586	565
dTPR2 (GH09432) GenBank submit	1488	1537

G C G C T G C A G C T G G A G A A G A C G C C G G A G A T T A A G C G A A T G C T G C G C G A G G C

AA950779	581	580
A1456556	274	273
A1457049	586	565
A1110008	679	678
AA201154	530	529
AA952015	783	782
A1516974	678	677
AA201429	367	386
AA392620	535	534
AA439655	570	569
A1544064	689	688
A1516760	532	531
AA978757	480	479
AA202799	368	367
AA735495	504	503
AA201527	371	370
AA438619	378	377
A1516274	682	691
AA440018	636	635
AA820977	589	588
A1455565	586	565
dTPR2 (GH09432) GenBank submit	1538	1587

C A A G T T T G C G T T G A A G A A G T C G A A G C G A A A G G A C T A C T A C A A G A T C C T G G

GenBank accession numbers

# FIGURE 11A (Continued)

AA950779	581		580
A1456556	274		273
A1457049	586		565
A1110008	679		678
AA201154	530		529
AA952015	783		782
A1516974	678		677
AA201429	367		386
AA392620	535		534
AA439655	570		569
A1544064	689		688
A1516760	532		531
AA978757	480		479
AA202799	368		367
AA735495	504		503
AA201527	371		370
AA438619	378		377
A1516274	692		691
AA440018	636		635
AA820977	589		588
A1455565	586		565
dTPR2 (GH09432) GenBank submit	1588	GCATTGGACGCAATGCGTCCGACGACGAGATCAAGAAGGCGTATCGCAA	1637

AA950779	581		580
A1456556	274		273
A1457049	586		565
A1110008	679		678
AA201154	530		529
AA952015	783		782
A1516974	678		677
AA201429	367		386
AA392620	535		534
AA439655	570		569
A1544064	689		688
A1516760	532		531
AA978757	480		479
AA202799	368		367
AA735495	504		503
AA201527	371		370
AA438619	378		377
A1516274	692		691
AA440018	636		635
AA820977	589		588
A1455565	586		565
dTPR2 (GH09432) GenBank submit	1638	AAGGCGCTGGTACATCATCCGGATCGACACGCAAAACAGCAGTGCCGAGGA	1687

AA950779	581		580
A1456556	274		273
A1457049	586		565
A1110008	679		678
AA201154	530		529
AA952015	783		782
A1516974	678		677
AA201429	367		386
AA392620	535		534
AA439655	570		569
A1544064	689		688
A1516760	532		531
AA978757	480		479
AA202799	368		367
AA735495	504		503
AA201527	371		370
AA438619	378		377
A1516274	692		691
AA440018	636		635
AA820977	589		588
A1455565	586		565
dTPR2 (GH09432) GenBank submit	1688	GCGCAAGGAGGAGGAGCTCAAGTTCAAGGAGGTGGGCGAGGCGTACGCCA	1737

AA950779	581		580
A1456556	274		273
A1457049	586		565
A1110008	679		678
AA201154	530		529
AA952015	783		782
A1516974	678		677
AA201429	367		386
AA392620	535		534
AA439655	570		569
A1544064	689		688
A1516760	532		531
AA978757	480		479
AA202799	368		367
AA735495	504		503
AA201527	371		370
AA438619	378		377
A1516274	692		691
AA440018	636		635
AA820977	589		588
A1455565	586		565
dTPR2 (GH09432) GenBank submit	1738	TACTGTGGATGCTCACAAAGAAGTCGCGCTACGACAGCGGCCAGGATATC	1787

# FIGURE 11A (Continued)

AA950779	581	580
AI456556	274	273
AI457049	566	565
AI110008	679	678
AA201154	530	529
AA952015	783	782
AI516974	678	677
AA201429	367	366
AA392620	535	534
AA439655	570	569
AI544064	689	688
AI516760	532	531
AA978757	480	479
AA202799	368	367
AA735495	504	503
AA201527	371	370
AA438619	378	377
AI516274	682	681
AA440018	636	635
AA820977	589	588
AI455565	585	565
dTPR2 (GH09432) GenBank submit	1788	1837
G A G G A G C A G G A G C A A G C C G A C T T C G A T C C G A A T C A A A T G T T C C G C A C A T T		
AA950779	581	580
AI456556	274	273
AI457049	566	565
AI110008	679	678
AA201154	530	529
AA952015	783	782
AI516974	678	677
AA201429	367	366
AA392620	535	534
AA439655	570	569
AI544064	689	688
AI516760	532	531
AA978757	480	479
AA202799	368	367
AA735495	504	503
AA201527	371	370
AA438619	378	377
AI516274	682	681
AA440018	636	635
AA820977	589	588
AI455565	585	565
dTPR2 (GH09432) GenBank submit	1838	1887
C T T C C A A T T C A A C G G C G G T G G C C G G A A T A A T T C A T C G T T C A A C T T T G A G T		
AA950779	581	580
AI456556	274	273
AI457049	566	565
AI110008	679	678
AA201154	530	529
AA952015	783	782
AI516974	678	677
AA201429	367	366
AA392620	535	534
AA439655	570	569
AI544064	689	688
AI516760	532	531
AA978757	480	479
AA202799	368	367
AA735495	504	503
AA201527	371	370
AA438619	378	377
AI516274	682	681
AA440018	636	635
AA820977	589	588
AI455565	585	565
dTPR2 (GH09432) GenBank submit	1888	1937
T C T A G G A T C C C A A C G A G T G T T G T T C A C C A C C A C A G A G A A G A A G A C C A T C T		
AA950779	581	580
AI456556	274	273
AI457049	566	565
AI110008	679	678
AA201154	530	529
AA952015	783	782
AI516974	678	677
AA201429	367	366
AA392620	535	534
AA439655	570	569
AI544064	689	688
AI516760	532	531
AA978757	480	479
AA202799	368	367
AA735495	504	503
AA201527	371	370
AA438619	378	377
AI516274	682	681
AA440018	636	635
AA820977	589	588
AI455565	585	565
dTPR2 (GH09432) GenBank submit	1938	1987
C A A T C C C A T A C T T T C T G C C T C A T C C G A A A C C A A C A T A C A G C A G C G C A C A A		

20250303 14:00:00

# FIGURE 11A (Continued)

AA950779	581	580
AI456556	274	273
AI457049	586	585
AI110008	679	678
AA201154	530	529
AA952015	783	782
AI516974	678	677
AA201429	387	386
AA392620	535	534
AA439655	570	569
AI544064	689	688
AI516760	532	531
AA978757	480	479
AA202799	388	387
AA735495	504	503
AA201527	371	370
AA438619	378	377
AI516274	692	691
AA440018	636	635
AA820977	589	588
AI455565	586	585
dTPR2 (GH09432) GenBank submit	1988	2037

ATTTTGAACCTCTTTTACATATTTCTTTTCCAAAAAGCAAGAAAAATACCAC

AA950779	581	580
AI456556	274	273
AI457049	586	585
AI110008	679	678
AA201154	530	529
AA952015	783	782
AI516974	678	677
AA201429	387	386
AA392620	535	534
AA439655	570	569
AI544064	689	688
AI516760	532	531
AA978757	480	479
AA202799	388	387
AA735495	504	503
AA201527	371	370
AA438619	378	377
AI516274	692	691
AA440018	636	635
AA820977	589	588
AI455565	586	585
dTPR2 (GH09432) GenBank submit	2238	2087

ATTTTGATTATGTTAACGAATGAATATATGCCAAGTTATTTGAAAAAATA

AA950779	581	580
AI456556	274	273
AI457049	586	585
AI110008	679	678
AA201154	530	529
AA952015	783	782
AI516974	678	677
AA201429	387	386
AA392620	535	534
AA439655	570	569
AI544064	689	688
AI516760	532	531
AA978757	480	479
AA202799	388	387
AA735495	504	503
AA201527	371	370
AA438619	378	377
AI516274	692	691
AA440018	636	635
AA820977	589	588
AI455565	586	585
dTPR2 (GH09432) GenBank submit	2088	2137

TTCTAAATCAAAATAATGCAACTAAATTTCCAGTGTAAGTTCACATTTT

AA950779	581	580
AI456556	274	273
AI457049	586	585
AI110008	679	678
AA201154	530	529
AA952015	783	782
AI516974	678	677
AA201429	387	386
AA392620	535	534
AA439655	570	569
AI544064	689	688
AI516760	532	531
AA978757	480	479
AA202799	388	387
AA735495	504	503
AA201527	371	370
AA438619	378	377
AI516274	692	691
AA440018	636	635
AA820977	589	588
AI455565	586	585
dTPR2 (GH09432) GenBank submit	2138	2187

AAATGTTCTTTCTTGGAATTTTTTTTTTCGGCAACATTAATAAATCATGGGA

[illegible][illegible]

AA950779	591	590
AI456556	274	273
AI457049	595	595
AI110008	679	678
AA201154	500	529
AA952015	783	782
AI516974	678	677
AA201429	357	395
AA392620	535	534
AA439655	570	569
AI544064	699	698
AI516760	532	531
AA978757	480	479
AA202799	368	367
AA735495	504	503
AA201527	371	370
AA438619	378	377
AI516274	682	681
AA440018	636	635
AA820977	599	598
AI5552	595	595
dTPP2 (GH09432) GenBank submit	2236	A A 2239

## ClustalW Formatted Alignments

A1456052  
 AA391402  
 A1514849  
 A1519786  
 A1546378  
 AA438290  
 AA820520  
 AA941597  
 A1455870  
 AA950277  
 AA941028  
 AA951077  
 AA951083  
 AA949900  
 AA541065  
 AA441093  
 dTPR2 (GH09432) GenBank submit



# FIGURE 11B (Continued)

AI456052  
AA391402  
AI514849  
AI519786  
AI546378  
AA438290  
AA820520  
AA941597  
AI455870  
AA950277  
AA941028  
AA951077  
AA951083  
AA949900  
AA541065  
AA441093  
dTPR2 (GH09432) GenBank submit

```

289 GCGCAAAAGTTGCGGCAATTTGGAGGCGACCAATTCAGGCGAACTACGATAC 332
319 GCGCAAAAGTTGCGGCAATTTGGAGGCGACCAATTCAGGCGAACTACGATAC 367
400 GCGCAAAAGTTGCGGCAATTTGGAGGCGACCAATTCAGGCGAACTACGATAC 440
389 GCGCAAAAGTTGCGGCAATTTGGAGGCGACCAATTCAGGCGAACTACGATAC 448
292 ATTTTGTAAATTGCCACCGCATGGACGACGAAGTAATTGAAATTAGCGACAG 341
301 ATTTTGTAAATTGCCACCGCATGGACGACGAAGTAATTGAAATTAGCGACAG 351
301 ATTTTGTAAATTGCCACCGCATGGACGACGAAGTAATTGAAATTAGCGACAG 353
301 ATTTTGTAAATTGCCACCGCATGGACGACGAAGTAATTGAAATTAGCGACAG 359
301 ATTTTGTAAATTGCCACCGCATGGACGACGAAGTAATTGAAATTAGCGACAG 359
301 ATTTTGTAAATTGCCACCGCATGGACGACGAAGTAATTGAAATTAGCGACAG 359
302 ATTTTGTAAATTGCCACCGCATGGACGACGAAGTAATTGAAATTAGCGACAG 359
305 ATTTTGTAAATTGCCACCGCATGGACGACGAAGTAATTGAAATTAGCGACAG 354
247
171
317 ATTTTGTAAATTGCCACCGCATGGACGACGAAGTAATTGAAATTAGCGACAG 363
317 ATTTTGTAAATTGCCACCGCATGGACGACGAAGTAATTGAAATTAGCGACAG 366
318 ATTTTGTAAATTGCCACCGCATGGACGACGAAGTAATTGAAATTAGCGACAG 367
348 ATTTTGTAAATTGCCACCGCATGGACGACGAAGTAATTGAAATTAGCGACAG 367

```

AI456052  
AA391402  
AI514849  
AI519786  
AI546378  
AA438290  
AA820520  
AA941597  
AI455870  
AA950277  
AA941028  
AA951077  
AA951083  
AA949900  
AA541065  
AA441093  
dTPR2 (GH09432) GenBank submit

```

339 GAAATCCTATCGCAATGTGGTCTTCTATTTTGGATAGTGCCCTTGAAATTTGG 389
368 GAAATCCTATCGCAATGTGGTCTTCTATTTTGGATAGTGCCCTTGAAATTTGG 417
450 GAAATCCTATCGCAATGTGGTCTTCTATTTTGGATAGTGCCCTTGAAATTTGG 498
449 GAAATCCTATCGCAATGTGGTCTTCTATTTTGGATAGTGCCCTTGAAATTTGG 498
342 CGAA--CCTGAAAGAAACCTCATCGAACTCCGAAATGGATGTGGAAATAA 386
351 CA--CCTGAAAGAAACCTCATCGAACTCCGAAATGGATGTGGAAATAA 386
351 GA--CCTGAAAGAAACCTCATCGAACTCCGAAATGGATGTGGAAATAA 396
351 GAA--CCTGAAAGAAACCTCATCGAACTCCGAAATGGATGTGGAAATAA 397
351 CG--CCTGAAAGAAACCTCATCGAACTCCGAAATGGATGTGGAAATAA 397
352 GA--CCTGAAAGAAACCTCATCGAACTCCGAAATGGATGTGGAAATAA 397
365 CA--CCTGAAAGAAACCTCATCGAACTCCGAAATGGATGTGGAAATAA 400
247
171
367 AT--CCTGAAAGAAACCTCATCGAACTCCGAAATGGATGTGGAAATAA 410
367 GA--CCTGAAAGAAACCTCATCGAACTCCGAAATGGATGTGGAAATAA 410
368 GA--CCTGAAAGAAACCTCATCGAACTCCGAAATGGATGTGGAAATAA 411
388 CG--CCTGAAAGAAACCTCATCGAACTCCGAAATGGATGTGGAAATAA 441

```

AI456052  
AA391402  
AI514849  
AI519786  
AI546378  
AA438290  
AA820520  
AA941597  
AI455870  
AA950277  
AA941028  
AA951077  
AA951083  
AA949900  
AA541065  
AA441093  
dTPR2 (GH09432) GenBank submit

```

389 CGCCCTCTGTTTGAAATATCGTCTAC TCAAGGGCTTGAGTGCCCTTGCAATTT 438
418 CGCCCTCTGTTTGAAATATCGTCTAC TCAAGGGCTTGAGTGCCCTTGCAATTT 467
500 CGCCCTCTGTTTGAAATATCGTCTAC TCAAGGGCTTGAGTGCCCTTGCAATTT 546
499 CGCCCTCTGTTTGAAATATCGTCTAC TCAAGGGCTTGAGTGCCCTTGCAATTT 548
389 CGACAGAACAGCC--AACCATCGATG--TCAAAGCAGAGCAAATTTGTGCC 435
395 CGACAGAACAGCC--AACCATCGATG--TCAAAGCAGAGCAAATTTGTGCC 442
397 CGACAGAACAGCC--AACCATCGATG--TCAAAGCAGAGCAAATTTGTGCC 443
398 CGACAGAACAGCC--AACCATCGATG--TCAAAGCAGAGCAAATTTGTGCC 444
398 CGACAGAACAGCC--AACCATCGATG--TCAAAGCAGAGCAAATTTGTGCC 444
398 CGACAGAACAGCC--AACCATCGATG--TCAAAGCAGAGCAAATTTGTGCC 444
401 CGACAGAACAGCC--AACCATCGATG--TCAAAGCAGAGCAAATTTGTGCC 447
247
171
413 CGACAGAACAGCC--AACCATCGATG--TCAAAGCAGAGCAAATTTGTGCC 456
411 CGACAGAACAGCC--AACCATCGATG--TCAAAGCAGAGCAAATTTGTGCC 457
412 CGACAGAACAGCC--AACCATCGATG--TCAAAGCAGAGCAAATTTGTGCC 458
445 CGACAGAACAGCC--AACCATCGATG--TCAAAGCAGAGCAAATTTGTGCC 491

```

AI456052  
AA391402  
AI514849  
AI519786  
AI546378  
AA438290  
AA820520  
AA941597  
AI455870  
AA950277  
AA941028  
AA951077  
AA951083  
AA949900  
AA541065  
AA441093  
dTPR2 (GH09432) GenBank submit

```

439 TTGGGCGCATGTGATGAGGCTTTGGACATTGCGGCTCAGTGTAATGAAACT 489
468 TTGGGCGCATGTGATGAGGCTTTGGACATTGCGGCTCAGTGTAATGAAACT 517
550 TTGGGCGCATGTGATGAGGCTTTGGACATTGCGGCTCAGTGTAATGAAACT 566
549 TTGGGCGCATGTGATGAGGCTTTGGACATTGCGGCTCAGTGTAATGAAACT 561
436 A-AGGACCGCGCAACCATTTGCCGAGGAGAGAGAGAACTGGGCAACGACC 484
443 A-AGGACCGCGCGCAACCATTTGCCGAGGAGAGAGAGAACTGGGCAACGACC 491
444 A-AGGACCGCGCGCAACCATTTGCCGAGGAGAGAGAGAACTGGGCAACGACC 492
445 A-AGGACCGCGCGCAACCATTTGCCGAGGAGAGAGAGAACTGGGCAACGACC 493
445 A-AGGACCGCGCGCAACCATTTGCCGAGGAGAGAGAGAACTGGGCAACGACC 493
445 A-AGGACCGCGCGCAACCATTTGCCGAGGAGAGAGAGAACTGGGCAACGACC 493
448 A-AGGACCGCGCGCAACCATTTGCCGAGGAGAGAGAGAACTGGGCAACGACC 497
247
171
460 A-AGGACCGCGCGCAACCATTTGCCGAGGAGAGAGAGAACTGGGCAACGACC 505
458 A-AGGACCGCGCGCAACCATTTGCCGAGGAGAGAGAGAACTGGGCAACGACC 506
459 A-AGGACCGCGCGCAACCATTTGCCGAGGAGAGAGAGAACTGGGCAACGACC 507
462 A-AGGACCGCGCGCAACCATTTGCCGAGGAGAGAGAGAACTGGGCAACGACC 540

```

GenBank submit

# FIGURE 11B (Continued)

AI456052	489	GGATACCACAT	CGG	CGGATG	CGAT	ATAC	GTGAG	AGG	TCT	GTG	CCTG	TACT	526	
AA391402	518	GGATACCACAT	CGG	CGGATG	CGAT	ATAC	GTGAG	AGG	TCT	GTG	CCTG	TACT	567	
AI514849	600	GGATACCACAT	CGG	CGGATG	CGAT	ATAC	GTGA						621	
AI519786	592												561	
AI546378	485	AATACAAAGGCGCAGAACTATCAGAATGCACTCAAGCTCTACACGGATGCC												554
AA438290	492	AATACAAAGGCGCAGAACTATCAGAATGCACTCAAGCTCTACACGGATGCC												541
AA820520	493	AATACAAAGGCGCAGAACTATCAGAATGCACTCAAGCTCTACACGGATGCC												542
AA941597	494	AATACAAAGGCGCAGAACTATCAGAATGCACTCAAGCTCTACACGGATGCC												543
AI455870	494	AATACAAAGGCGCAGAACTATCAGAATGCACTCAAGCTCTACACGGATGCC												543
AA950277	494	AATACAAAGGCGCAGAACTATCAGAATGCACTCAAGCTCTACACGGATGCC												543
AA941028	488													487
AA951077	247													246
AA951083	171													170
AA949900	509	AAT												511
AA541065	507	AATACAAAGGCGCAGAACTATCAGAATGCACTCAAGCTCTACACGGATGCC												556
AA441093	508	AATACAAAGGCGCAGAACTATCAGAATGCACTCAAGCTCTACACGGATGCC												557
dTPR2 (GH09432) GenBank submit	541	AATACAAAGGCGCAGAACTATCAGAATGCACTCAAGCTCTACACGGATGCC												592

AI456052	539	ACACGG	ACAACTGGACAAGGGAATTCTTTCATTTTCGAGCGCGCCCTGACC	586
AA391402	558	ACACGG	ACAACTGGACAAGGGAATTCTTTCATTTTCGAGCGCGCCCTGACC	617
AI514849	632			631
AI519786	592			561
AI546378	535	ATA	TCC	541
AA438290	542	ATA	TCC	560
AA820520	543	ATA	TCC	562
AA941597	544	ATA	TCC	562
AI455870	544	ATA	TCC	559
AA950277	544	ATA	TCC	559
AA941028	488			487
AA951077	247			246
AA951083	171			170
AA949900	512			511
AA541065	557	ATA	TCC	600
AA441093	558	ATA	TCC	566
dTPR2 (GH09432) GenBank submit	591	ATA	TCC	623

AI456052	599	C		585
AA391402	618	CTCGACCCGGGACC	ACTACCAGTCCAAGCAGATGCGC	657
AI514849	632			631
AI519786	562			561
AI546378	542			541
AA438290	561			560
AA820520	593	CTACATGATGCTGCTCAACTATAAT		617
AA941597	583			582
AI455870	551			550
AA950277	547			546
AA941028	488			487
AA951077	247			246
AA951083	171			170
AA949900	512			511
AA541065	607	CTACATGATGCTGCTCAACTATAATAGCGCCCTGACCGACGCCCGACACG		655
AA441093	597			596
dTPR2 (GH09432) GenBank submit	641	CTACATGATGCTGCTCAACTATAATAGCGCCCTGACCGACGCCCGACACG		691

AI456052	590			585
AA391402	654			653
AI514849	632			631
AI519786	562			561
AI546378	542			541
AA438290	561			560
AA820520	618			617
AA941597	583			582
AI455870	551			550
AA950277	547			546
AA941028	488			487
AA951077	247			246
AA951083	171			170
AA949900	512			511
AA541065	657	CCATACGCATCGATCCGGGCTTCGAGAAGGCCTACGTCCGTGT		656
AA441093	597			596
dTPR2 (GH09432) GenBank submit	691	CCATACGCATCGATCCGGGCTTCGAGAAGGCCTACGTCCGTGTGGCCAAAG		743

GenBank accession numbers

# FIGURE 11B (Continued)

AI456052	590	590
AA391402	654	653
AI514849	632	631
AI519786	562	561
AI546378	542	541
AA438290	561	560
AA820520	618	617
AA941597	583	582
AI455870	551	550
AA950277	547	546
AA941028	488	487
AA951077	247	246
AA951083	171	170
AA949900	512	511
AA541065	700	699
AA441093	597	596
dTPR2 (GH09432) GenBank submit	741	740

T G C T G T C T G G C C C T G G G C G A C A T T A T T G G C A C C G A A C A G G C C G T C A A A A T

AI456052	590	590
AA391402	654	653
AI514849	632	631
AI519786	562	561
AI546378	542	541
AA438290	561	560
AA820520	618	617
AA941597	583	582
AI455870	551	550
AA950277	547	546
AA941028	488	487
AA951077	247	246
AA951083	171	170
AA949900	512	511
AA541065	700	699
AA441093	597	596
dTPR2 (GH09432) GenBank submit	791	790

G G T C A A C G A G C T G A A T T C G C T T A G C A C G G C T G T T G C T G C C G A A C A G A C G G

AI456052	590	590
AA391402	654	653
AI514849	632	631
AI519786	562	561
AI546378	542	541
AA438290	561	560
AA820520	618	617
AA941597	583	582
AI455870	551	550
AA950277	547	546
AA941028	488	487
AA951077	247	246
AA951083	171	170
AA949900	512	511
AA541065	700	699
AA441093	597	596
dTPR2 (GH09432) GenBank submit	841	840

C G G C G C A A A A G T T G C G C C A A T T G G A G G C C A C C A T T C A G G C G A A C T A C G A T

AI456052	590	590
AA391402	654	653
AI514849	632	631
AI519786	562	561
AI546378	542	541
AA438290	561	560
AA820520	618	617
AA941597	583	582
AI455870	551	550
AA950277	547	546
AA941028	488	487
AA951077	247	246
AA951083	171	170
AA949900	512	511
AA541065	700	699
AA441093	597	596
dTPR2 (GH09432) GenBank submit	891	890

A C G A A A T C C T A T C G C A A T G T G G T C T T C T A T T T G G A T A G T G C C T T G A A A T T

bioRxiv preprint doi: <https://doi.org/10.1101/202209>; this version posted September 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

# FIGURE 11B (Continued)

AI456052	590	585
AA391402	654	653
AI514849	632	631
AI519786	562	561
AI546378	542	541
AA438290	561	560
AA820520	618	617
AA941597	583	582
AI455870	551	550
AA950277	547	546
AA941028	488	487
AA951077	247	246
AA951083	171	170
AA949900	512	511
AA541065	700	699
AA441093	597	596
dTPR2 (GH09432) GenBank submit	941	893

G G C G C C C G C C T G T T T G A A A T A T C G T C T A C T C A A G G C T G A G T G C C T T G C A T

AI456052	590	585
AA391402	654	653
AI514849	632	631
AI519786	562	561
AI546378	542	541
AA438290	561	560
AA820520	618	617
AA941597	583	582
AI455870	551	550
AA950277	547	546
AA941028	488	487
AA951077	247	246
AA951083	171	170
AA949900	512	511
AA541065	700	699
AA441093	597	596
dTPR2 (GH09432) GenBank submit	991	1039

T T T T G G G G C G A T G T G A T G A G G C C T T G G A C A T T G C G G T C A G T G T A A T G A A A

AI456052	590	585
AA391402	654	653
AI514849	632	631
AI519786	562	561
AI546378	542	541
AA438290	561	560
AA820520	618	617
AA941597	583	582
AI455870	551	550
AA950277	547	546
AA941028	488	487
AA951077	247	246
AA951083	171	170
AA949900	512	511
AA541065	700	699
AA441093	597	596
dTPR2 (GH09432) GenBank submit	1041	1094

C T G G A T A C C A C A T C G G C G G A T G C G A T A T A C G T G A G A G G T C T G T G C C T G T A

AI456052	590	585
AA391402	654	653
AI514849	632	631
AI519786	562	561
AI546378	542	541
AA438290	561	560
AA820520	618	617
AA941597	583	582
AI455870	551	550
AA950277	547	546
AA941028	488	487
AA951077	247	246
AA951083	171	170
AA949900	512	511
AA541065	700	699
AA441093	597	596
dTPR2 (GH09432) GenBank submit	1091	1140

C T A C A C G G A C A A C C T G G A C A A G G G A A T T C T T C A T T T C G A G C G C G C C C T G A

001100 20362860

# FIGURE 11B (Continued)

AI456052	590		590
AA391402	654		653
AI514849	632		631
AI519786	562		561
AI546378	542		541
AA438290	561		560
AA820520	618		617
AA941597	583		582
AI455870	551		550
AA950277	547		546
AA941028	488		487
AA951077	247		246
AA951083	171		170
AA949900	512		511
AA541065	700		699
AA441093	597		596
dTPR2 (GH09432) GenBank submit	1141	CCCTCGACCCGGACCACTACAAGTCCAAGCAGATGCGCAGCAAATGCAAG	1139

AI456052	590		589
AA391402	654		653
AI514849	632		631
AI519786	562		561
AI546378	542		541
AA438290	561		560
AA820520	618		617
AA941597	583		582
AI455870	551		550
AA950277	547		546
AA941028	488		487
AA951077	247		246
AA951083	171		170
AA949900	512		511
AA541065	700		699
AA441093	597		596
dTPR2 (GH09432) GenBank submit	1191	CAGCTCAAGGAGATGAAGGAGAACGGCAATATGCTATTCAAGTCGGGTCTG	1189

AI456052	590		589
AA391402	654		653
AI514849	632		631
AI519786	562		561
AI546378	542		541
AA438290	561		560
AA820520	618		617
AA941597	583		582
AI455870	551		550
AA950277	547		546
AA941028	488		487
AA951077	247		246
AA951083	171		170
AA949900	512		511
AA541065	700		699
AA441093	597		596
dTPR2 (GH09432) GenBank submit	1241	GTATCGCGAGGCACACGTTATCTACACGGACGCCCTGAAGATCGATGAAC	1239

AI456052	590		589
AA391402	654		653
AI514849	632		631
AI519786	562		561
AI546378	542		541
AA438290	561		560
AA820520	618		617
AA941597	583		582
AI455870	551		550
AA950277	547		546
AA941028	488		487
AA951077	247		246
AA951083	171		170
AA949900	512		511
AA541065	700		699
AA441093	597		596
dTPR2 (GH09432) GenBank submit	1291	ACAACAAGGATATCAATTTCGAAATTGCTTTACAATCGGGCTTTGGTCAAC	1289

GenBank accession numbers

# FIGURE 11B (Continued)

AI456052	590	589
AA391402	654	653
AI514849	632	631
AI519786	562	561
AI546378	542	541
AA438290	561	560
AA820520	618	617
AA941597	583	582
AI455870	551	550
AA950277	547	546
AA941028	488	487
AA951077	247	246
AA951083	171	170
AA949900	512	511
AA541065	700	699
AA441093	597	596
dTPR2 (GH09432) GenBank submitt	1341	1340

ACGCGTATTGGCAATTTGCGAGAGGCCGTGGCCGATTGCAATCGAGTGC T

AI456052	590	589
AA391402	654	653
AI514849	632	631
AI519786	562	561
AI546378	542	541
AA438290	561	560
AA820520	618	617
AA941597	583	582
AI455870	551	550
AA950277	547	546
AA941028	488	487
AA951077	247	246
AA951083	171	170
AA949900	512	511
AA541065	700	699
AA441093	597	596
dTPR2 (GH09432) GenBank submitt	1391	1390

GGAGCTGAATAGTCAGTATCTGAAGGCTCTGTTGCTGCGAGCGCGCTGC T

AI456052	590	589
AA391402	654	653
AI514849	632	631
AI519786	562	561
AI546378	542	541
AA438290	561	560
AA820520	618	617
AA941597	583	582
AI455870	551	550
AA950277	547	546
AA941028	488	487
AA951077	247	246
AA951083	171	170
AA949900	512	511
AA541065	700	699
AA441093	597	596
dTPR2 (GH09432) GenBank submitt	1441	1440

ACAAATGATCTGGAGAAGTTTCGAGGAGTCGGTGGCGGACTATGAGACGGCG

AI456052	590	589
AA391402	654	653
AI514849	632	631
AI519786	562	561
AI546378	542	541
AA438290	561	560
AA820520	618	617
AA941597	583	582
AI455870	551	550
AA950277	547	546
AA941028	488	487
AA951077	247	246
AA951083	171	170
AA949900	512	511
AA541065	700	699
AA441093	597	596
dTPR2 (GH09432) GenBank submitt	1491	1490

CTGCAGCTGGAGAAGACGCCGGAGATTAAAGCGAATGCTGCGCGAGGCCAA

# FIGURE 11B (Continued)

AI456052	590	589
AA391402	654	653
AI514849	632	631
AI519786	562	561
AI546378	542	541
AA438290	561	560
AA820520	618	617
AA941597	583	582
AI455870	551	550
AA950277	547	546
AA941028	488	487
AA951077	247	246
AA951083	171	170
AA949900	512	511
AA541065	700	699
AA441093	597	596
dTPR2 (GH09432) GenBank submit	1541	1540

G T T T G C G T T G A A G A A G T C G A A G C G A A A G G A C T A C T A C A A G A T C C T G G G C A

AI456052	590	589
AA391402	654	653
AI514849	632	631
AI519786	562	561
AI546378	542	541
AA438290	561	560
AA820520	618	617
AA941597	583	582
AI455870	551	550
AA950277	547	546
AA941028	488	487
AA951077	247	246
AA951083	171	170
AA949900	512	511
AA541065	700	699
AA441093	597	596
dTPR2 (GH09432) GenBank submit	1591	1590

T T G G A C G C A A T G C G T C C G A C G A C G A G A T C A A G A A G G C G T A T C G C A A A A A G

AI456052	590	589
AA391402	654	653
AI514849	632	631
AI519786	562	561
AI546378	542	541
AA438290	561	560
AA820520	618	617
AA941597	583	582
AI455870	551	550
AA950277	547	546
AA941028	488	487
AA951077	247	246
AA951083	171	170
AA949900	512	511
AA541065	700	699
AA441093	597	596
dTPR2 (GH09432) GenBank submit	1641	1640

G C G C T G G T A C A T C A T C C G G A T C G A C A C G C A A A C A G C A G T G C C G A G G A G C G

AI456052	590	589
AA391402	654	653
AI514849	632	631
AI519786	562	561
AI546378	542	541
AA438290	561	560
AA820520	618	617
AA941597	583	582
AI455870	551	550
AA950277	547	546
AA941028	488	487
AA951077	247	246
AA951083	171	170
AA949900	512	511
AA541065	700	699
AA441093	597	596
dTPR2 (GH09432) GenBank submit	1691	1690

C A A G G A G G A G G A G C T C A A G T T C A A G G A G G T G G G C G A G G C G T A C G C C A T A C

# FIGURE 11B (Continued)

AI456052	590		589
AA391402	654		653
AI514849	632		631
AI519786	562		561
AI546378	542		541
AA438290	561		560
AA820520	618		617
AA941597	583		582
AI455870	551		550
AA950277	547		546
AA941028	488		487
AA951077	247		246
AA951083	171		170
AA949900	512		511
AA541065	700		699
AA441093	597		596
dTPR2 (GH09432) GenBank submitt	1741	T G T C G G A T G C T C A C A A G A A G T C G C G C T A C G A C A G C G G C C A G G A T A T C G A G	1740

AI456052	590		589
AA391402	654		653
AI514849	632		631
AI519786	562		561
AI546378	542		541
AA438290	561		560
AA820520	618		617
AA941597	583		582
AI455870	551		550
AA950277	547		546
AA941028	488		487
AA951077	247		246
AA951083	171		170
AA949900	512		511
AA541065	700		699
AA441093	597		596
dTPR2 (GH09432) GenBank submitt	1791	G A G C A G G A G C A A G C C G A C T T C G A T C C G A A T C A A A T G T T C C G C A C A T T C T T	1790

AI456052	590		589
AA391402	654		653
AI514849	632		631
AI519786	562		561
AI546378	542		541
AA438290	561		560
AA820520	618		617
AA941597	583		582
AI455870	551		550
AA950277	547		546
AA941028	488		487
AA951077	247		246
AA951083	171		170
AA949900	512		511
AA541065	700		699
AA441093	597		596
dTPR2 (GH09432) GenBank submitt	1841	C C A A T T C A A C G G C G G T G G C C G G A A T A A T T C A T C G T T C A A C T T T G A G T T C T	1840

AI456052	590		589
AA391402	654		653
AI514849	632		631
AI519786	562		561
AI546378	542		541
AA438290	561		560
AA820520	618		617
AA941597	583		582
AI455870	551		550
AA950277	547		546
AA941028	488		487
AA951077	247		246
AA951083	171		170
AA949900	512		511
AA541065	700		699
AA441093	597		596
dTPR2 (GH09432) GenBank submitt	1891	A G G A T C C C A A C G A G T G T T G T T C A C C A C C A C A G A G A A G A A G A C C A T C T C A A	1890

GenBank accession numbers

# FIGURE 11B (Continued)

AI456052	590		589
AA391402	654		653
AI514849	632		631
AI519786	562		561
AI546378	542		541
AA438290	561		560
AA820520	618		617
AA941597	583		582
AI455870	551		550
AA950277	547		546
AA941028	488		487
AA951077	247		246
AA951083	171		170
AA949900	512		511
AA541065	700		699
AA441093	597		596
dTPR2 (GH09432) GenBank submit	1941	T C C C A T A C T T T T C T G C C T C A T C C G A A C C A A C A T A C A G C A G C G C A C A A A T T	1991

AI456052	590		589
AA391402	654		653
AI514849	632		631
AI519786	562		561
AI546378	542		541
AA438290	561		560
AA820520	618		617
AA941597	583		582
AI455870	551		550
AA950277	547		546
AA941028	488		487
AA951077	247		246
AA951083	171		170
AA949900	512		511
AA541065	700		699
AA441093	597		596
dTPR2 (GH09432) GenBank submit	1991	T T G A A C T C T T T T T A C A T A T T T C T T T T T C C A A A A A G C A A G A A A A T A C C A C A T T	2041

AI456052	590		589
AA391402	654		653
AI514849	632		631
AI519786	562		561
AI546378	542		541
AA438290	561		560
AA820520	618		617
AA941597	583		582
AI455870	551		550
AA950277	547		546
AA941028	488		487
AA951077	247		246
AA951083	171		170
AA949900	512		511
AA541065	700		699
AA441093	597		596
dTPR2 (GH09432) GenBank submit	2041	T T G A T T A T G T T A A C G A A T G A A T A T A T G C C A A G T T A T T T G A A A A A A T A T T C	2091

AI456052	590		589
AA391402	654		653
AI514849	632		631
AI519786	562		561
AI546378	542		541
AA438290	561		560
AA820520	618		617
AA941597	583		582
AI455870	551		550
AA950277	547		546
AA941028	488		487
AA951077	247		246
AA951083	171		170
AA949900	512		511
AA541065	700		699
AA441093	597		596
dTPR2 (GH09432) GenBank submit	2091	T A A A T C A A A A T A A T G C A A C T A A A T T T C C A G T G T A A G T T C A C A T T T T T A A A	2141

bioRxiv preprint doi: <https://doi.org/10.1101/202203>; this version posted March 2, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FIGURE 11B (Continued)**

Accession	Position	Sequence
AI456052	590	
AA391402	654	
AI514849	632	
AI519786	562	
AI546378	542	
AA438290	561	
AA820520	618	
AA941597	593	
AI455870	551	
AA950277	547	
AA941028	488	
AA951077	247	
AA951083	171	
AA949900	512	
AA541065	700	
AA441093	597	
dTPR2 (GH09432) GenBank submit	2141	T G T T C T T T C T T G G A T T T T T T T T T C G G C A A C A T T A A T A A A T C A T G G G A G A T

[illegible]

## ClustalW Formatted Alignments



# FIGURE 12 (Continued)

AA950937	122	TGATGAACAGCTTTATGCCCGATCCCTTCATGCAGGTCTCGCCCTTTGAC	171
AI113626	140	TGATGAACAGCTTTATGCCCGATCCCTTCATGCAGGTCTCGCCCTTTGAC	189
AA978719	154	TGATGAACAGCTTTATGCCCGATCCCTTCATGCAGGTCTCGCCCTTTGAC	203
AI260802	159	TGATGAACAGCTTTATGCCCGATCCCTTCATGCAGGTCTCGCCCTTTGAC	208
AI541593	161	TGATGAACAGCTTTATGCCCGATCCCTTCATGCAGGTCTCGCCCTTTGAC	210
AI541599	161	TGATGAACAGCTTTATGCCCGATCCCTTCATGCAGGTCTCGCCCTTTGAC	210
AA695052	161	TGATGAACAGCTTTATGCCCGATCCCTTCATGCAGGTCTCGCCCTTTGAC	210
AI107772	161	TGATGAACAGCTTTATGCCCGATCCCTTCATGCAGGTCTCGCCCTTTGAC	210
AI113621	161	TGATGAACAGCTTTATGCCCGATCCCTTCATGCAGGTCTCGCCCTTTGAC	210
AI515548	180	TGATGAACAGCTTTATGCCCGATCCCTTCATGCAGGTCTCGCCCTTTGAC	229
AI532198	186	TGATGAACAGCTTTATGCCCGATCCCTTCATGCAGGTCTCGCCCTTTGAC	235
AI389024	189	TGATGAACAGCTTTATGCCCGATCCCTTCATGCAGGTCTCGCCCTTTGAC	238
AA391063	193	TGATGAACAGCTTTATGCCCGATCCCTTCATGCAGGTCTCGCCCTTTGAC	242
AA940727	195	TGATGAACAGCTTTATGCCCGATCCCTTCATGCAGGTCTCGCCCTTTGAC	244
AA802232	196	TGATGAACAGCTTTATGCCCGATCCCTTCATGCAGGTCTCGCCCTTTGAC	245
AI531980	197	TGATGAACAGCTTTATGCCCGATCCCTTCATGCAGGTCTCGCCCTTTGAC	246
dmf cDNA only	198	TGATGAACAGCTTTATGCCCGATCCCTTCATGCAGGTCTCGCCCTTTGAC	247
AA264488	201	TGATGAACAGCTTTATGCCCGATCCCTTCATGCAGGTCTCGCCCTTTGAC	250

AA950937	172	CAGGGATTCCAGCA	GAACGGCTCTCATGGAGCGTCCGCAGATGCCGGCCA	220
AI113626	190	CAGGGATTCCAGCA	GAACGGCTCTCATGGAGCGTCCGCAGATGCCGGCCA	238
AA978719	204	CAGGGATTCCAGCA	GAACGGCTCTCATGGAGCGTCCGCAGATGCCGGCCA	252
AI260802	209	CAGGGATTCCAGCA	GAACGGCTCTCATGGAGCGTCCGCAGATGCCGGCCA	257
AI541593	211	CAGGGATTCCAGCA	GAACGGCTCTCATGGAGCGTCCGCAGATGCCGGCCA	259
AI541599	211	CAGGGATTCCAGCA	GAACGGCTCTCATGGAGCGTCCGCAGATGCCGGCCA	259
AA695052	211	CAGGGATTCCAGCA	GAACGGCTCTCATGGAGCGTCCGCAGATGCCGGCCA	259
AI107772	211	CAGGGATTCCAGCA	GAACGGCTCTCATGGAGCGTCCGCAGATGCCGGCCA	259
AI113621	211	CAGGGATTCCAGCA	GAACGGCTCTCATGGAGCGTCCGCAGATGCCGGCCA	259
AI515548	230	CAGGGATTCCAGCA	GAACGGCTCTCATGGAGCGTCCGCAGATGCCGGCCA	278
AI532198	236	CAGGGATTCCAGCA	GAACGGCTCTCATGGAGCGTCCGCAGATGCCGGCCA	284
AI389024	239	CAGGGATTCCAGCA	GAACGGCTCTCATGGAGCGTCCGCAGATGCCGGCCA	287
AA391063	243	CAGGGATTCCAGCA	GAACGGCTCTCATGGAGCGTCCGCAGATGCCGGCCA	292
AA940727	245	CAGGGATTCCAGCA	GAACGGCTCTCATGGAGCGTCCGCAGATGCCGGCCA	293
AA802232	246	CAGGGATTCCAGCA	GAACGGCTCTCATGGAGCGTCCGCAGATGCCGGCCA	294
AI531980	247	CAGGGATTCCAGCA	GAACGGCTCTCATGGAGCGTCCGCAGATGCCGGCCA	295
dmf cDNA only	248	CAGGGATTCCAGCA	GAACGGCTCTCATGGAGCGTCCGCAGATGCCGGCCA	296
AA264488	251	CAGGGATTCCAGCA	GAACGGCTCTCATGGAGCGTCCGCAGATGCCGGCCA	298

AA950937	221	TGCCAGCCATGGGCTCTTCGGCATGCCCA	TGATGCCCA	CA	AACTTTAA	268
AI113626	239	TGCCAGCCATGGGCTCTTCGGCATGCCCA	TGATGCCCA		AACTTTAA	284
AA978719	253	TGCCAGCCATGGGCTCTTCGGCATGCCCA	TGATGCCCA		AACTTTAA	298
AI260802	258	TGCCAGCCATGGGCTCTTCGGCATGCCCA	TGATGCCCA		AACTTTAA	303
AI541593	260	TGCCAGCCATGGGCTCTTCGGCATGCCCA	TGATGCCCA		AACTTTAA	305
AI541599	260	TGCCAGCCATGGGCTCTTCGGCATGCCCA	TGATGCCCA		AACTTTAA	305
AA695052	260	TGCCAGCCATGGGCTCTTCGGCATGCCCA	TGATGCCCA		AACTTTAA	305
AI107772	260	TGCCAGCCATGGGCTCTTCGGCATGCCCA	TGATGCCCA		AACTTTAA	305
AI113621	260	TGCCAGCCATGGGCTCTTCGGCATGCCCA	TGATGCCCA		AACTTTAA	305
AI515548	279	TGCCAGCCATGGGCTCTTCGGCATGCCCA	TGATGCCCA		AACTTTAA	324
AI532198	285	TGCCAGCCATGGGCTCTTCGGCATGCCCA	TGATGCCCA		AACTTTAA	330
AI389024	288	TGCCAGCCATGGGCTCTTCGGCATGCCCA	TGATGCCCA		AACTTTAA	333
AA391063	293	TGCCAGCCATGGGCTCTTCGGCATGCCCA	TGATGCCCA		AACTTTAA	335
AA940727	294	TGCCAGCCATGGGCTCTTCGGCATGCCCA	TGATGCCCA		AACTTTAA	339
AA802232	295	TGCCAGCCATGGGCTCTTCGGCATGCCCA	TGATGCCCA		AACTTTAA	340
AI531980	296	TGCCAGCCATGGGCTCTTCGGCATGCCCA	TGATGCCCA		AACTTTAA	341
dmf cDNA only	297	TGCCAGCCATGGGCTCTTCGGCATGCCCA	TGATGCCCA		AACTTTAA	342
AA264488	299	TGCCAGCCATGGGCTCTTCGGCATGCCCA	TGATGCCCA		AACTTTAA	343

AA950937	299	TCGCCCTGTTGAACGCTGATATTGGTGGCAATTCAGGGCGCATCCTTCTGCC	318
AI113626	295	TCGCCCTGTTGAACGCTGATATTGGTGGCAATTCAGGGCGCATCCTTCTGCC	334
AA978719	299	TCGCCCTGTTGAACGCTGATATTGGTGGCAATTCAGGGCGCATCCTTCTGCC	348
AI260802	304	TCGCCCTGTTGAACGCTGATATTGGTGGCAATTCAGGGCGCATCCTTCTGCC	353
AI541593	306	TCGCCCTGTTGAACGCTGATATTGGTGGCAATTCAGGGCGCATCCTTCTGCC	355
AI541599	306	TCGCCCTGTTGAACGCTGATATTGGTGGCAATTCAGGGCGCATCCTTCTGCC	355
AA695052	306	TCGCCCTGTTGAACGCTGATATTGGTGGCAATTCAGGGCGCATCCTTCTGCC	355
AI107772	306	TCGCCCTGTTGAACGCTGATATTGGTGGCAATTCAGGGCGCATCCTTCTGCC	355
AI113621	306	TCGCCCTGTTGAACGCTGATATTGGTGGCAATTCAGGGCGCATCCTTCTGCC	355
AI515548	325	TCGCCCTGTTGAACGCTGATATTGGTGGCAATTCAGGGCGCATCCTTCTGCC	374
AI532198	331	TCGCCCTGTTGAACGCTGATATTGGTGGCAATTCAGGGCGCATCCTTCTGCC	380
AI389024	334	TCGCCCTGTTGAACGCTGATATTGGTGGCAATTCAGGGCGCATCCTTCTGCC	383
AA391063	336	TCGCCCTGTTGAACGCTGATATTGGTGGCAATTCAGGGCGCATCCTTCTGCC	385
AA940727	340	TCGCCCTGTTGAACGCTGATATTGGTGGCAATTCAGGGCGCATCCTTCTGCC	389
AA802232	341	TCGCCCTGTTGAACGCTGATATTGGTGGCAATTCAGGGCGCATCCTTCTGCC	390
AI531980	342	TCGCCCTGTTGAACGCTGATATTGGTGGCAATTCAGGGCGCATCCTTCTGCC	391
dmf cDNA only	343	TCGCCCTGTTGAACGCTGATATTGGTGGCAATTCAGGGCGCATCCTTCTGCC	392
AA264488	344	TCGCCCTGTTGAACGCTGATATTGGTGGCAATTCAGGGCGCATCCTTCTGCC	393

[illegible]

AA950937	469	G	G	T	C	A	T	C	A	C	A	T	C	G	G	C	G	A	G	C	G	G	G	C	A	C	A	C	A	T	T	A	T	T	G	A	G	A	A	G	A	G	C	A	G	G	A	C	A	T	517
AI113626	484	G	G	T	C	A	T	C	A	C	A	T	C	G	G	C	G	A	G	C	G	G	G	C	A	C	A	C	A	T	T	A	T	T	G	A	G	A	A	G	A	G	C	A	G	G	A	C	A	T	533
AA978719	498	G	G	T	C	A	T	C	A	C	A	T	C	G	G	C	G	A	G	C	G	G	G	C	A	C	A	C	A	T	T	A	T	T	G	A	G	A	A	G	A	G	C	A	G	G	A	C	A	T	547
AI260802	503	G	G	T	C	A	T	C	A	C	A	T	C	G	G	C	G	A	G	C	G	G	G	C	A	C	A	C	A	T	T	A	T	T	G	A	G	A	A	G	A	G	C	A	G	G	A	C	A	T	552
AI541593	505	G	G	T	C	A	T	C	A	C	A	T	C	G	G	C	G	A	G	C	G	G	G	C	A	C	A	C	A	T	T	A	T	T	G	A	G	A	A	G	A	G	C	A	G	G	A	C	A	T	554
AI541599	505	G	G	T	C	A	T	C	A	C	A	T	C	G	G	C	G	A	G	C	G	G	G	C	A	C	A	C	A	T	T	A	T	T	G	A	G	A	A	G	A	G	C	A	G	G	A	C	A	T	554
AA695052	505	G	G	T	C	A	T	C	A	C	A	T	C	G	G	C	G	A	G	C	G	G	G	C	A	C	A	C	A	T	T	A	T	T	G	A	G	A	A	G	A	G	C	A	G	G	A	C	A	T	554
AI107772	505	G	G	T	C	A	T	C	A	C	A	T	C	G	G	C	G	A	G	C	G	G	G	C	A	C	A	C	A	T	T	A	T	T	G	A	G	A	A	G	A	G	C	A	G	G	A	C	A	T	554
AI113621	505	G	G	T	C	A	T	C	A	C	A	T	C	G	G	C	G	A	G	C	G	G	G	C	A	C	A	C	A	T	T	A	T	T	G	A	G	A	A	G	A	G	C	A	G	G	A	C	A	T	554
AI515548	524	G	G	T	C	A	T	C	A	C	A	T	C	G	G	C	G	A	G	C	G	G	G	C	A	C	A	C	A	T	T	A	T	T	G	A	G	A	A	G	A	G	C	A	G	G	A	C	A	T	573
AI532198	498	G	G	T	C	A	T	C	A	C	A	T	C	G	G	C	G	A	G	C	G	G	G	C	A	C	A	C	A	T	T	A	T	T	G	A	G	A	A	G	A	G	C	A	G	G	A	C	A	T	497
AI389024	533	G	G	T	C	A	T	C	A	C	A	T	C	G	G	C	G	A	G	C	G	G	G	C	A	C	A	C	A	T	T	A	T	T	G	A	G	A	A	G	A	G	C	A	G	G	A	C	A	T	582
AA391063	536	G	G	T	C	A	T	C	A	C	A	T	C	G	G	C	G	A	G	C	G	G	G	C	A	C	A	C	A	T	T	A	T	T	G	A	G	A	A	G	A	G	C	A	G	G	A	C	A	T	540
AA940727	539	G	G	T	C	A	T	C	A	C	A	T	C	G	G	C	G	A	G	C	G	G	G	C	A	C	A	C	A	T	T	A	T	T	G	A	G	A	A	G	A	G	C	A	G	G	A	C	A	T	588
AA802232	540	G	G	T	C	A	T	C	A	C	A	T	C	G	G	C	G	A																																	

# FIGURE 12 (Continued)

AA950937	518	GCGCTCAGGACAACCTGGAGGAGCGCCAGGAGTTCATTAAATCTGGAGGAGG	567
AI113626	534	GCGCTCAGGACAACCTGGAGGAGCGCCAGGAGTTCATTAAATCTGGAGGAGG	563
AA978719	548	GCGCTCAGGACAACCTGGAGGAGCGCCAGGAGTTCATTAAATCTGGAGGAGG	567
AI260802	553	GCGCTCAGGACAACCTGGAGGAGCGCCAGGAGTTCATTAAATCTGGAGGAGG	602
AI541593	555	GCGCTCAGGACAACCTGGAGGAGCGCCAGGAGTTCATTAAATCTGGAGGAGG	599
AI541599	555	GCGCTCAGGACAACCTGGAGGAGCGCCAGGAGTTCATTAAATCTGGAGGAGG	604
AA695052	555	GCGCTCAGGACAACCTGGAGGAGCGCCAGGAGTTCATTAAATCTGGAGGAGG	585
AI107772	555	GCGCTCAGGACAACCTGGAGGAGCGCCAGGAGTTCATTAAATCTGGAGGAGG	591
AI113621	555	GCGCTCAGGACAACCTGGAGGAGCGCCAGGAGTTCATTAAATCTGGAGGAGG	604
AI515548	574	GCGCTCAGGACAACCTGGAGGAGCGCCAGGAGTTCATTAAATCTGGAGGAGG	623
AI532198	498	GCGCTCAGGACAACCTGGAGGAGCGCCAGGAGTTCATTAAATCTGGAGGAGG	497
AI389024	583	GCGCTCAGGACAACCTGGAGGAGCGCCAGGAGTTCATTAAATCTGGAGGAGG	632
AA391063	541	GCGCTCAGGACAACCTGGAGGAGCGCCAGGAGTTCATTAAATCTGGAGGAGG	540
AA940727	589	GCGCTCAGGACAACCTGGAGGAGCGCCAGGAGTTCATTAAATCTGGAGGAGG	638
AA802232	590	GCGCTCAGGACAACCTGGAGGAGCGCCAGGAGTTCATTAAATCTGGAGGAGG	639
AI531980	591	GCGCTCAGGACAACCTGGAGGAGCGCCAGGAGTTCATTAAATCTGGAGGAGG	627
dmlf cDNA only	592	GCGCTCAGGACAACCTGGAGGAGCGCCAGGAGTTCATTAAATCTGGAGGAGG	641
AA264488	594	GCGCTCAGGACAACCTGGAGGAGCGCCAGGAGTTCATTAAATCTGGAGGAGG	643

AA950937	558	GAGAGAGCCGAGCAGTTTGACAGGAGTTTACATCGCGCGCTAGTCGCGGA	617
AI113626	564	GAGAGAGCCGAGCAGTTTGACAGGAGTTTACATCGCGCGCTAGTCGCGGA	563
AA978719	598	GAGAGAGCCGAGCAGTTTGACAGGAGTTTACATCGCGCGCTAGTCGCGGA	647
AI260802	603	GAGAGAGCCGAGCAGTTTGACAGGAGTTTACATCGCGCGCTAGTCGCGGA	652
AI541593	600		599
AI541599	605	GAGAG	606
AA695052	586		585
AI107772	592		591
AI113621	605	GAGAGAGCCGAGCAGTTTGACAGGAGTTTACATCGCGCGCTAGTCGCGGA	654
AI515548	624	GAGAGAGCCGAGCAGTTTGACAGGAGTTTACATCGCGCGCTAGTCGCGGA	673
AI532198	498		497
AI389024	633	GAGAGAGCCGAGCAGTTTGACAGGAGTTTACATCGCGCGCTAGTCGCGGA	682
AA391063	541		540
AA940727	639	GAGAGAGCCGAGCAGTTTGACAGGAGTTTACATCGCGCTAG	681
AA802232	640	AGAGAGCGAGCAGTTTGACAGGAGTTTACATCGCGCGCTAG	675
AI531980	628		627
dmlf cDNA only	642	GAGAGAGCCGAGCAGTTTGACAGGAGTTTACATCGCGCGCTAGTCGCGGA	691
AA264488	644	GAGAGAGCCGAGCAGTTTGACAGGAGTTTACATCGCGCGCTAG	686

AA950937	618	GCGGTGCAGTCAAGACATCATGCTGGTGGCATGCAGGCCATCATGCCCGC	666
AI113626	564		563
AA978719	648	GGCGTGCAGTCAAGACATCATGCTGGTGGCATGCAGGCCATCATGCCCGC	697
AI260802	653	GCGGTGCAGTCAAGACATCATGCTGGTGGCATGCAGGCCATCATGCCCGC	701
AI541593	600		599
AI541599	600		606
AA695052	586		585
AI107772	592		591
AI113621	655	GCG	657
AI515548	674	GCGGTGCAGTCAAGACATCATGCTGGTGGCATGCA-GCCATCATG	717
AI532198	498		497
AI389024	683	GCG	685
AA391063	541		540
AA940727	682		681
AA802232	676		675
AI531980	628		627
dmlf cDNA only	682	GCGGTGCAGTCAAGACATCATGCTGGTGGCATGCAGGCCATCATGCCCGC	741
AA264488	687		686

AA950937	667	C	667
AI113626	564		563
AA978719	698	CCGTCCAGCGGCACACACCTCGACGTTGACCATTGAGCCAGTGGAGGACG	747
AI260802	702	CCGTCCA	708
AI541593	600		599
AI541599	600		606
AA695052	586		585
AI107772	592		591
AI113621	658		657
AI515548	718		717
AI532198	498		497
AI389024	686		685
AA391063	541		540
AA940727	682		681
AA802232	676		675
AI531980	628		627
dmlf cDNA only	742	CCGTCCAGCGGCACACACCTCGACGTTGACCATTGAGCCAGTGGAGGACG	791
AA264488	687		686

CCGTCCAGCGGCACACACCTCGACGTTGACCATTGAGCCAGTGGAGGACG

# FIGURE 12 (Continued)

AA950937	668		667
AI113626	564		563
AA978719	748	ACGACGACGATGATGC	763
AI260802	709		708
AI541593	600		599
AI541599	609		608
AA695052	586		585
AI107772	592		591
AI113621	658		657
AI515548	718		717
AI532198	498		497
AI389024	686		685
AA391063	541		540
AA940727	682		681
AA802232	676		675
AI531980	628		627
dmf cDNA only	792	ACGACGACGATGATGATGACTGTGTAATCCAGGAGCAGCAACCGGTTTCGC	841
AA264488	687		686
AA950937	668		667
AI113626	564		563
AA978719	764		763
AI260802	709		708
AI541593	600		599
AI541599	609		608
AA695052	586		585
AI107772	592		591
AI113621	658		657
AI515548	718		717
AI532198	498		497
AI389024	686		685
AA391063	541		540
AA940727	682		681
AA802232	676		675
AI531980	628		627
dmf cDNA only	842	TCCTCCGCGGGCCGCCATTATTCCAGTGCGCCAACGGCACCGCAGAACAG	891
AA264488	687		686
AA950937	668		667
AI113626	564		563
AA978719	764		763
AI260802	709		708
AI541593	600		599
AI541599	609		608
AA695052	586		585
AI107772	592		591
AI113621	658		657
AI515548	718		717
AI532198	498		497
AI389024	686		685
AA391063	541		540
AA940727	682		681
AA802232	676		675
AI531980	628		627
dmf cDNA only	892	ATATAATTACTAAATCTAAAGTCAATACAGTATATTTTACTAACTATCCG	941
AA264488	687		686
AA950937	668		667
AI113626	564		563
AA978719	764		763
AI260802	709		708
AI541593	600		599
AI541599	609		608
AA695052	586		585
AI107772	592		591
AI113621	658		657
AI515548	718		717
AI532198	498		497
AI389024	686		685
AA391063	541		540
AA940727	682		681
AA802232	676		675
AI531980	628		627
dmf cDNA only	942	ATAAACAGAAACAGAAATTGCATACTATAAAATTTCTGCTAATTACATTCC	991
AA264488	687		686

# FIGURE 12 (Continued)

AA950937	668		667
AI113626	564		563
AA978719	764		763
AI260802	709		708
AI541593	600		599
AI541599	609		608
AA695052	586		585
AI107772	592		591
AI113621	658		657
AI515548	718		717
AI532198	498		497
AI389024	686		685
AA391063	541		540
AA940727	682		681
AA802232	676		675
AI531980	628		627
dmf cDNA only	992	C A A C T G C G T T C A A A C G A A A C G A A T A T C G A A T C G A A A T C A T A G A A T G C A C A	1041
AA264488	687		686

AA950937	668		667
AI113626	564		563
AA978719	764		763
AI260802	709		708
AI541593	600		599
AI541599	609		608
AA695052	586		585
AI107772	592		591
AI113621	658		657
AI515548	718		717
AI532198	498		497
AI389024	686		685
AA391063	541		540
AA940727	682		681
AA802232	676		675
AI531980	628		627
dmf cDNA only	1042	G A G C A G C A T A C A T C C A C A T C C C T A T G C C G C C A A T C C G A G G C G C C A A C A A C	1091
AA264488	687		686

AA950937	668		667
AI113626	564		563
AA978719	764		763
AI260802	709		708
AI541593	600		599
AI541599	609		608
AA695052	586		585
AI107772	592		591
AI113621	658		657
AI515548	718		717
AI532198	498		497
AI389024	686		685
AA391063	541		540
AA940727	682		681
AA802232	676		675
AI531980	628		627
dmf cDNA only	1082	G T G C C G T A A A A C A T T T T C A C A C G G A G G A C G A A G C G G C C A G C T C C T A C A A G	1141
AA264488	687		686

AA950937	668		667
AI113626	564		563
AA978719	764		763
AI260802	709		708
AI541593	600		599
AI541599	609		608
AA695052	586		585
AI107772	592		591
AI113621	658		657
AI515548	718		717
AI532198	498		497
AI389024	686		685
AA391063	541		540
AA940727	682		681
AA802232	676		675
AI531980	628		627
dmf cDNA only	1142	G C G G T C A A G C G C G G C A A G A A G A G T A G T A G A A A C G T G A T C A T C T G T A T G C	1191
AA264488	687		686

# FIGURE 12 (Continued)

AA950937	668		667
AI113626	564		563
AA978719	764		763
AI260802	709		708
AI541593	600		599
AI541599	609		608
AA695052	586		585
AI107772	592		591
AI113621	658		657
AI515548	718		717
AI532198	498		497
AI389024	686		685
AA391063	541		540
AA940727	682		681
AA802232	676		675
AI531980	628		627
dmf cDNA only	1192	C A A C A T C T T C C G C A T C G C A C A C T C A A A A C A C T A G G A A G C A A A G C G T T G G	1241
AA264488	687		686

AA950937	668		667
AI113626	564		563
AA978719	764		763
AI260802	709		708
AI541593	600		599
AI541599	609		608
AA695052	586		585
AI107772	592		591
AI113621	658		657
AI515548	718		717
AI532198	498		497
AI389024	686		685
AA391063	541		540
AA940727	682		681
AA802232	676		675
AI531980	628		627
dmf cDNA only	1242	G T T C T G T T C C A T A G C A G G A A A C C A A T T C A A A T A T T T T T A A C A A A C A C A	1291
AA264488	687		686

AA950937	668		667
AI113626	564		563
AA978719	764		763
AI260802	709		708
AI541593	600		599
AI541599	609		608
AA695052	586		585
AI107772	592		591
AI113621	658		657
AI515548	718		717
AI532198	498		497
AI389024	686		685
AA391063	541		540
AA940727	682		681
AA802232	676		675
AI531980	628		627
dmf cDNA only	1292	A T T C T T T A C C A G T T C T G T C T T A T C C T G C G T G A G T C G A C C A G A A T G C A A C A	1341
AA264488	687		686

AA950937	668		667
AI113626	564		563
AA978719	764		763
AI260802	709		708
AI541593	600		599
AI541599	609		608
AA695052	586		585
AI107772	592		591
AI113621	658		657
AI515548	718		717
AI532198	498		497
AI389024	686		685
AA391063	541		540
AA940727	682		681
AA802232	676		675
AI531980	628		627
dmf cDNA only	1342	C T A A A A A A T G T A C A A C T T C A A G A T G C T A T T G A T G T G C A C G C A G G A T A C A G	1391
AA264488	687		686

[illegible]

AA950937	668		667
A1113626	664		563
AA978719	764		763
A1260802	709		708
A1541593	600		599
A1541599	609		608
AA695052	585		585
A1107772	582		581
A1113621	658		657
A1515548	718		717
A1532198	498		497
A1389024	685		685
AA391063	541		540
AA940727	682		681
AA802232	676		675
A1531980	628		627
dmf1 cDNA only	1542	CTCAAGTTTTTAGATTTTGTAGCCACTAAGCTTTAAATTATGGATGCCA	1541
AA264488	687		686

# FIGURE 12 (Continued)

AA950937	668		667
AI113626	564		563
AA978719	764		763
AI260802	709		708
AI541593	600		599
AI541599	609		608
AA695052	596		585
AI107772	592		591
AI113621	658		657
AI515548	718		717
AI532198	498		497
AI389024	686		685
AA391063	541		540
AA940727	682		681
AA802232	676		675
AI531980	628		627
dmf cDNA only	1592	G T T A G C G T G C A A A T G A A C A C A A T T G A T T T G A A G G C T C C G A A C G A T A G A A A	1641
AA264488	687		686

AA950937	668		667
AI113626	564		563
AA978719	764		763
AI260802	709		708
AI541593	600		599
AI541599	609		608
AA695052	596		585
AI107772	592		591
AI113621	658		657
AI515548	718		717
AI532198	498		497
AI389024	686		685
AA391063	541		540
AA940727	682		681
AA802232	676		675
AI531980	628		627
dmf cDNA only	1642	A C A A C A A T T A C C A A T T C C C C A A A T A C A T G T A A T T C G T A A G G C C T A A G T A A	1691
AA264488	687		686

AA950937	668		667
AI113626	564		563
AA978719	764		763
AI260802	709		708
AI541593	600		599
AI541599	609		608
AA695052	596		585
AI107772	592		591
AI113621	658		657
AI515548	718		717
AI532198	498		497
AI389024	686		685
AA391063	541		540
AA940727	682		681
AA802232	676		675
AI531980	628		627
dmf cDNA only	1692	A T G T T A A C G T G A A T T T A A T T A A A T G G T A A T T A C A T T A T A A T A G T A A A A A A	1741
AA264488	687		686

AA950937	668	667
AI113626	564	563
AA978719	764	763
AI260802	709	708
AI541593	600	599
AI541599	609	608
AA695052	596	585
AI107772	592	591
AI113621	658	657
AI515548	718	717
AI532198	498	497
AI389024	686	685
AA391063	541	540
AA940727	682	681
AA802232	676	675
AI531980	628	627
dmf cDNA only	1742	A A A A A A A A A A A A A A A
AA264488	687	686

## FIGURE 13A

hTPR2 Protein 484 amino acids

MAATEPELLDDQEAKREAETFKEQGNAYYAKKDYNEAYNYYTKAIDMCPKNA  
SYYGNRAATLMMLGRFREALGDAQSVRLDDSFVRGHLREGKCHLSLGNAMA  
ACRSFQRALELDHKNAQAQQEFKNANAVMEYEKIAETDFEKRDFRKVVFCMDR  
ALEFAPACHRFKILKAECAMLGRYPEAQSVASDILRMDSTNADALYVRGLCLY  
YEDCIEKAVQFFVQALRMAPDHEKACIACRNAKALKAKKEDGNKAFKEGNYKL  
AYELYTEALGIDPNNIKTNAKLYCNRGTVNSKLRKLDDAIEDCTNAVKLDDTYI  
KAYLRRAQCYMDTEQYEEAVRDYEVYQTEKTEHKQLLKNAQLELKSKSRK  
DYYKILGVDKNASEDEIKKAYRKRALMHPDRHSGASAEVQKEEEKKFKEVGE  
AFTILSDPKKKTRYDSGQDLDEEGMNMGDFDPNNIFKAFFGGPGGFSFEASGPGN  
FFFQFG

## FIGURE 13B

hTPR2 cDNA 1756 base pairs

CGGCTGCCGCGGAGTGCGATGTGGTAATGGCGGCGACCGAGCCGGAGCTGCT  
CGACGACCAAGAGGCGAAGAGGGAAGCAGAGACTTTCAAGGAACAAGGAAA  
TGCATACTATGCCAAGAAAGATTACAATGAAGCTTATAATTATTATACAAAA  
GCCATAGATATGTGTCCTAAAAATGCTAGCTATTATGGTAATCGAGCAGCCA  
CCTTGATGATGCTTGGAAGGTTCCGGGAAGCTCTTGGAGATGCACAACAGTC  
AGTGAGGTTGGATGACAGTTTTGTCCGGGGACATCTACGAGAGGGCAAGTGC  
CACCTCTCTCTGGGGAATGCCATGGCAGCATGTCGCAGCTTCCAGAGAGCCC  
TAGAACTGGATCATAAAAAATGCTCAGGCACAACAAGAGTTCAAGAATGCTAA  
TGCAGTCATGGAATATGAGAAAATAGCAGAAACAGATTTTGAGAAGCGAGA  
TTTTTCGGAAGGTTGTTTTCTGCATGGACCGTGCCCTAGAATTTGCCCTGCCT  
GCCATCGCTTCAAATCCTCAAGGCAGAATGTTTAGCAATGCTGGGTCGTTAT  
CCGGAAGCACAGTCTGTGGCTAGTGACATTCTACGAATGGATTCCACCAATG  
CAGATGCTCTGTATGTACGAGGTCTTTGCCTTTATTACGAAGATTGTATTGAG  
AAGGCAGTTCAGTTTTTCGTACAGGCTCTCAGGATGGCTCCTGACCACGAGA  
AGGCCTGCATTGCCTGCAGAAATGCCAAAGCACTCAAAGCAAAGAAAGAAG  
ATGGGAATAAAGCATTAAAGGAAGGAAATTACAACTAGCATATGAACTGTA  
CACAGAAGCCCTGGGGATAGACCCCAACAATATAAAAACAAATGCTAACTC  
TACTGTAATCGGGGTACGGTTAATTCCAAGCTTAGGAACTAGATGATGCAA  
TAGAAGACTGCACAAATGCAGTGAAGCTTGATGACACTTACATAAAAGCCTA  
CTTGAGAAGAGCTCAGTGTTACATGGACACAGAACAGTATGAAGAAGCAGTA  
CGAGACTATGAAAAAGTATACCAGACAGAGAAAACAAAAGAACACAAACAG  
CTCCTAAAAAATGCGCAGCTGGAAGTGAAGAAGAGTAAGAGGAAAGATTAC  
TACAAGATTCTAGGAGTGGACAAGAATGCCTCTGAGGACGAGATCAAGAAA  
GCTTATCGGAAACGGGCCTTGATGCACCATCCAGATCGGCATAGTGGAGCCA  
GTGCTGAGGTTCAGAAGGAGGAGGAGAAGAAGTTCAAGGAAGTTGGAGAGG  
CCTTTACTATCCTCTCTGATCCCAAGAAAAAGACTCGCTATGACAGTGGACAG  
GACCTAGATGAGGAGGGCATGAATATGGGTGATTTTGATCCAAACAATATCT  
TCAAGGCATTCTTTGGCGGTCTGGCGGCTTCAGCTTTGAAGCATCTGGTCCA  
GGGAATTTCTTTTTTCAATTTGGCTAATGAAGGGCAACCACCCAGAACCCAG  
AAAATGCAGATTCAGTCTTAACTTTGAATGTGGAAACAGTTCACCTCCTC  
CCTTCATCACGTCTCCGTGTGCTTAGAGCAGTTTCGTTTTCTCAGTTGGATGCC  
CTGTGTCTCTGTGAGTGGGGTGGAGCAAAGGGAACCAATGCCGAAGACCGAG  
GGCAGGGGAGGGAGGCGGGGGTGGACAGGGAGGCAGCTTGTGAATTTTTGT  
TTTACTGTTTAACTTTATTAAAAAAGAAAAAAAAAAAAAAAAA

## FIGURE 14 A

hMLF Protein 268 amino acids

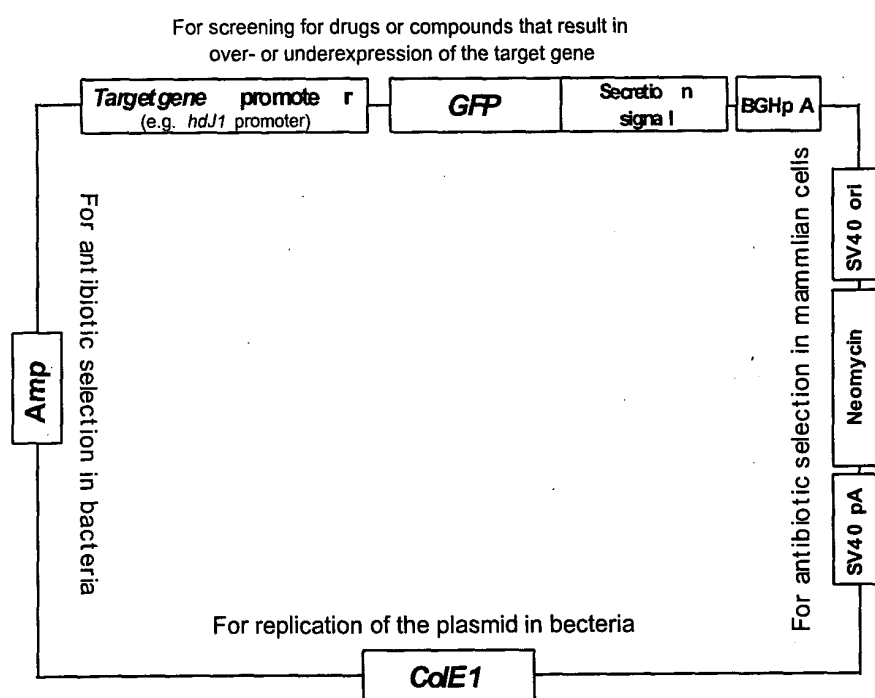
MFRMLNSSFEDDPFFSEILAHRENMRQMIRSFSEPFGRDLLSISDGRGRAHNRRG  
HNDGEDSLTHTDVSSFQTMQMVSNMRNYMQKLERNFGQLSVDPNGHSFCSSS  
VMTYSKIGDEPPKVFQASTQTRRAPGGIKETRKAMRDSDSGLEKMAIGHHHDR  
AHVIKKSKNKKTGDEEVNQEFINMNESDAHAFDEEWQSEVLKYKPGRHNLGNT  
RMRSVGHENPGSRELKRREKPQQSPAIEHGRRSNVLGDKLHIKGSSSVKSNKK

## FIGURE 14B

hMLF cDNA 1116 base pairs

GTTATGTGTTCCCGTCCGTACTGGAGGCTAGCTCTTGTCGCGGCCGCGGCGAG  
TTAACATCGTTTTTCCAATCTGTCCGCGGCTGCCGCCACCCAAGACAGAGCCA  
GAATGTTCAAGGATGCTGAACAGCAGTTTTGAGGATGACCCCTTCTTCTCTGAG  
TCCATTCTTGCACACCGAGAAAATATGCGACAGATGATAAGAAGTTTTTCTG  
AACCCCTTTGGAAGAGACTTGCTCAGTATCTCTGATGGTAGAGGGAGAGCTCA  
TAATCGTAGAGGACATAATGATGGTGAAGATTCTTTGACTCATACAGATGTC  
AGCTCTTCCAGACCATGGACCAAATGGTGTCAAATATGAGAACTATATGC  
AGAAATTAGAAAGAAACTTCGGTCAACTTTTCAAGTGGATCCAAATGGACATTC  
ATTTTGTCTTCTCCTCAGTTATGACTTATTCCAAAATAGGAGATGAACCGCCAA  
AGGTTTTTTCAGGCCTCAACTCAAACTCGTCGAGCTCCAGGAGGAATAAAGGA  
AACCAGGAAAGCAATGAGAGATTCTGACAGTGGACTAGAAAAAATGGCTAT  
TGGTCATCATATCCATGACCGAGCTCATGTCATTAAAAAGTCAAAGAACAAG  
AAGACTGGAGATGAAGAGGTCAACCAGGAGTTCATCAATATGAATGAAAGC  
GATGCTCATGCTTTTGATGAGGAGTGGCAAAGTGAGGTTTTGAAGTACAAAC  
CAGGACGACACAATCTAGGAAACACTAGAATGAGAAGTGTGGCCATGAGA  
ATCCTGGCTCCCGAGAACTTAAAAGAAGGGAGAAACCTCAACAAAGTCCAGC  
CATTGAACATGGAAGGAGATCAAATGTTTTGGGGGACAAACTCCACATCAAA  
GGCTCATCTGTGAAAAGCAACAAAAAATAAATAGCCATGCATTTGATTTGTT  
TAGTTTTGATTGTTTTAACAGTTAGTAATGGTGCTGGGTAATAAGCATAAGAC  
CAATCTCTTGCTGTAAATCAGTTCTGTCCTTGGCAACTTTCTTCTGATATCTG  
AATGTTTCATGAAGGTCCTAGCTTTATATTGTCCCTCTTTTAGGAATAAAATTTT  
GATTTTCAACAAAAAAA

**FIGURE 15**



## FIGURE 16

dHDJ1 5' region, 24333 base pairs

TTACGGTTTATTTACTATTACTCTAGTTAATCAAATAAACTGTATAATTCCTGG  
CTTGTACAATAATTTTGCTAACACGCCGATGCGTTCGATCTTTTTTTTTTACCGC  
TCTCCGTCGTATTCATCATGGTACATATTACATCCAACATACTTTATTTTTTT  
GGGTATTAAACATTGGCAATATCGCTGCTCGCCGCCGTTTCGGTTATGCTCTAT  
AAATAAAAGGGGGGGCGCCGCTAAAATTATAATAAAATTTTCATGGGTCCTAA  
ATCTAGTCTCGAAATCTATGTACAAAGTTTGCTTGCATGCTGGTTAGGCATAG  
GTTCTTAACGTATTATTGGGTTGCTTTATTTCCATTCTGCGCAGTTGTGCAGCC  
TGTTTAGTGTTTGCTTTACGGGGTTAACATTTTTTAAAAATGAAACATTAGA  
GCGGTAACCTTGTTGTCTGATTATTGGCGTCATTAAAGCGGTATCGCCAGCAC  
GCGATTGATGCAAGGATACCGATTCAATGAAATAAAAACGAATTCAGCCAAA  
CACAATCTTTCATTTCTTTTTTTTTATCGTACTTAATGATAGCCTTAGTTTCTA  
ATGGGACTGTGTGCTTCGGTGAAGGTTGGGGATGATTTTGGGAGGCAACAAT  
TATGTTCTAGCTTATAGCTTACAGTCCTACGCCTACTCCTATTTCTAATATGTT  
CATCATCAGCAGTTAAAAACGTTTACAAAACCTCATGCGAAATTGAAATCCA  
ATAACAAATGCACACGCCGCAGTCGCATCGGCGTCATCTCTTCTCCTGACCC  
TCGCCTATCCGCATCCAGTTAGGTTTGCTGCTGCTGCTGCGCCGACGGTTGTC  
GCCGACTGAAGCCACCGCCGGCGGACAGATGTCGTTGCAGGGCTCGCTGCTG  
CTGGAACCTTGGCGCTGCCTGGTCTCCGAAGCGGTTGAACTTGAACCTGTTGC  
GCTGCTGGAAGTTCTGGCGATAGTTCTGATTGTAGAATCGCGGAAATCCTCCA  
CCTCCGCCGTTCTTGTTCCAGCGCTTCTGGCCCTCGTACTCCTGGAATGGATT  
GTACCCGGGCGTGCTGTTGCTGGCATTGTTTCCCTTAGCCGAACCGGACTTCA  
CCTTCCGCTGACGTCCACGATCCATCTCGTTCTCTTCGTCGTCGTCGATGTCCC  
GCTGCCGCTGCTCACGCGCATCCACCAGTAGCTACGGAACAGAAATATCAA  
GCATTAGGCTAGAGTTCGGACCTTGTGAATGGGGAGGCTTGGCTGGCTGGCT  
GACGCATGCGCTAGTTAATGGAGCTTATGCAGATGAGTACGGTCGCTCGCGA  
ACAAGCACTGGGAATATGCACATTGTATTCGAAATGGGTGAGTGGCTTACGG  
TTCACGGTTCAGTGAACAGGTTATCAGGCAAAACGGTAACGGCACAACGGT  
TGAATTTATGGCGTATCAGGCGGTTGAAATGAAAGAAACAACGTGCCGGCCA  
GCAGTCAAATCATAAGCTTCATTGCACGGGAAACGGATGCGGAGTCATCGG  
GTGAATTACCTAGGCTCCGGTGCAGTCACTCTCTCCCGCAATGACTTTTGCAA  
CTCTCTCTACACTTTTACGCTCGCTGAACGGAGGACGCGTTGTGGTGACCGC  
CCGGTTGGGAACGGATACCAGCAACGCAGCCATCACAGACTATTCGGGGTAA  
TCGTATTATTTGTATTTGTTTTGTGTGGTATGTGCTTAGTGGGGAAAAAGAAG  
AAGCGTCGCTCTGCCGCCGACGCTTCTACCTCCTACCGGCCGTCCGTGAGAC  
GATCCGGATCGGGTGCGTCAGCGGTCTGTCTGTTACCGCCACTGCAATTACG  
ACCACATCTTTACTGTCACTGCCACTAGTCACTGCCGCGTCGACTGCAACCGA  
GCCCTCGACGATATCGCTGCCTTCCACACTGCCGTGACCAGCTATCCGTTTCG  
CACAAACCAACTCAAAGTCTAAATGAATGGGGATAATGTGGAAACAAATG  
CAAATTACAAACAAGTTCGTTTAGTAAATCAACTCAATCGAATTGCATTTTAT  
GCAACAGCTAAGCGAACGACATAGAAAACAAAAAAGAAGACCAAAGAGCCA  
GTAAATAATAAAGAATTAGTTAAACCCGCAAAAAGAGAACCAATTTATGTA  
CATTTTCATCGTATTAAGCCCGCAACTTGTTATTTTTGAAGCACAGACCCAAA  
GAAAGTGTTAACCATGCATAGATTTAGTATCTACGTAGTGACATGGTCACA  
AGGGATAGATAAGCGCTTCAAGGTGAATGCCTCTCTAAACTCACCTCCTTTTC

FIGURE 16 (Continued)

GAGCTCCGCGGGCTTGCCATTCCAACCTGAGCACGGGGGAGCCGTATCCACGA  
TACGATTGCTTCAGCAGCTCATTGATGGTGCTCCCATTCGAGGTGGCATTGCT  
CTGGTAGCCATTGCCACCCTTGGCTGCTGGTGCGACTTGAGCGCACTGCCGT  
TGAGCAACTTTTGGCGCTTGGCGCCTGGCGTTGCGGGCGAGTCCGTGGAGGG  
CTTGCTACTCGAGAAGGGATTGCGATGGTTCTTGTTTCGGCGTCGGAATTTTCA  
CCGGCGATCCCTCCACCACCACGACGTCAACATCATCTTCGATGGCATCGAC  
CTCATCGTTACGCGTAACTTTCCAGATACCCGTTTTTCGATTTGATGACCGCTG  
GCGAGGGTGGCGTCTTTGGCGATGGTGGTGCTTTGCTTTGAGACTGTGATTGC  
TTCTGTGGGTGCCAGCCATTCGTTAGCTGAATGCTGGGCTCCTCCTCATCGTC  
GTCCTCATCACTGTCTGCGGACTTTTTTAGGCTCTTGAATATCTCGTCTATGGC  
ATCAGTCTTTTGTGTTGCTGTTGTTAACGTGAACCGATGACGAGGCTGAGCCGT  
TGGTGTGGCTGCCATTGGTCTTACTGTGACCATTGGTCTGGCCAGACTCTTGG  
TCCGACTCGCTAGAATCCTCCCCACTTGGACGCTTGCAGGGGATTCCGGTAGAG  
GCGCCTCCTCCTCCTCAGATGCTGATTTCGTAGGGCACAAGACTTTTCAGTGGC  
GTTTTGACGGGAGTCTTTACTTGAATCTTTACAGGAGACTTTGCCTTAGGCTC  
CGTATGATTCTCCGTCATATTGGGCATGCTCGGTAGCTGGGCTGTCGTTGGTC  
TGGGCTTCATCTCATCCTCGATGTCTTCATCATCCGAGGAAATTGGCAGATAT  
TGTTGCTGGTTTTATTGATTTATGGTTGCTGCTGCTGTTGCTGCTTGGGGAGGA  
ACTTTTGTACCATTTCAGTGGGAGCTGAAGTAGCTTCGCCTTTAGCGTGAG  
CGCCAACCAAGTGGAGGCTTCGCAGTGTCTGAACTTTCCGGTACCCAGCTG  
GAGTCCGTTCTGTGGACTTTGCTGGTTTTGTTGCTTGAAGTGGATGGCGGTTTT  
CTGGGCATTTCCATTAGTGTAACCGTTTGTCTCACCCGCCGGCAGTTGAGGAC  
CAATGAAACGCGTTGGTGAAGGCGAAGACACAGTCGCCGCCGGCACTGGCGT  
TGTGCTGTGTCCGTTGGTCAAACGCACTCCATTGGGCCTGTTGGCCGCCGGAC  
TGGCAGCCTGTGAGAGGTCCAGTTCGAAAAACATTATATAGGCATTTGTGTT  
GCACACACTGTGCATTGCGATTGGCCGCACGTAGCTGTCTCGTCAAGTTGTAA  
AAGCTGCCCGTATCCGTGGAGCCAATGGCCGTGTAGTGACCGCAGTGCTGGG  
ACGCCCCCAAGTGAGTGACCATCGACACCAGGCGATAGGTGAGCGGTTGAGC  
CTGAGCTGCTTGTGAACGGGCTGCGTATTTGCTCAAATCTATGCGTGACTTGA  
AGGAAATCTGCTTGGTCAAGTTTGTGTCGATCATGGAGAATCGCTTCAGCTGT  
ATACAAAGCGTGATTGGGGCACGCTCCAAAGAGAATTGCTTTGTGGCAGATA  
CCTGCAAGCGATACGTTTAAATAAAATGAACTACAGAACAAAGGTCACAAAG  
ACCTACCTTCTTCTTGCATCCCTCGCACTTGTAGCCCATATCCTCTAGCCGTT  
GCGAGAAAAGTGTCCCTCGAAAGCATCCTCCAAGGAGTCTGCCTTGCGGATG  
TCGAGCAACAGATCCTGGAAGTGCTGAAACGTAATGGACACATGGTTGCAGC  
TCAGACAGCGCACCTCGCTGCGCAGATAGCCGCCAAAGATCTGTCCCAGCGG  
CGTGGTCTCCTTAACCAACTGATCCAGCTCTTTGTAGTTACGAAACCGCATCA  
AATACGCCCGCTCCATGGCCTCGACCAGGAAGCGCAGGAAGTCTGCGCATC  
CTCTTGGCGACCAACGACCATGTGTTTGCAGATCTGCTTTAGCTTCGAGTAGA  
TGAGGAAGGGTCTGACGGCCGACTGATTGCTTTGGGTGGCCAAAAGTGT  
GGTCATGGCGCAAATGATGCAACCGCTGCCAGGTTTCGGCCACATTGCAGTCA  
GCCAGATGCGCCTGCTCCGAAACGAGCCAATTGGCCAGGGCGGGTATGTGCA  
GGAGCGCCTGAAGCGTTGAGTTGAGGTAGCAGGTGTTGCCACATTGATCAT  
GCCCGTGCCACCTGCCATTTGCGCTCCGACTGCTTCCAGCCAATGCGTATGT  
TCTCCCGCGGATAGAGGACCCTCTTCGGCTTGGGCAGCTCATTGGGATTGCTT  
GTCGGATGCTGATGATTGTGGTGGTTGATGTGGTGCGACTGATTGTTCCGGGTG



## FIGURE 16 (Continued)

TCGTGCGCGGCCAGGCATTTTAGAACGAAAAGTCTGCTTTCGAAAATAATGG  
CAATTCCTTCCCTCGTGTCTTCTCCGACTGCGGATTCTCTTTTCGCTTCATTTTC  
GTCATTTGGGGATGCCAACTCGCGAGTGGCCAAGTGACGCGATAGGCCTCTC  
GAAATGTCCTAAAGCATTTACGATATTTACAAAAATGTATTTTCGATGTTTTT  
TTAACAATAAAAAATTGGTTTAAATTTAATAAGACATTTGTTACCTTGAATAT  
GTAAGCAATATCTTATTGAAAGGCTTGCAGCGACATTTTTTATTTATGCCTAC  
TATTCAAGTTATAAATTTAATTTTTATAACGGTATTTTTTACACCTTATCAGCAC  
ATATCGATAAGTGTGATTGGGAACGACAACCCATCGGCACAATGTTGATGCA  
ATTGTTGAGCTAGCCTTCATAATTAGTCGCAATCAATCGAGCAGAATGGCTTC  
ATCCACAGGTCTCCTGGTGGTGTCCAACATCAAGCACCTTGGCAAATCCCTGC  
GAGCCATCGAGAAGTACGTGAATTCAGTGTACATCCACCTAAATGTGGCGGG  
GTCAACGTCCACGACGTACCAAGTTCCACCGCCTCCGGTTTGGGGTTCGTCTAA  
TCTCGCAGCTGTACGCCAACAGCAGCAGCTATGTGGGCAAGCAGTTGGACCT  
TCGCGTCCTTGTCTCTCCCCTACGACCAGGTGCCAATGGATCCCTGAAGTTGC  
GCCAGCCCGTCGACCTAATCTTCTCGGATGCACATCATCCGGAGCTGTGCGAC  
AGGCTTCGCGCGGATCTTAACATCAGCAAGCCAACAATCTTCCTGGATGACT  
CGGTCATCTCGGATTTAAGTGCCGAGCAGGATGACACCCAGCCGCCTAAGGT  
GTATCCCTCGGTTGTCTGGGCGGAACATTCGATCGCATCCATCTGGGACACA  
AGATATTCCTCACCCAGGCTGTACTGCGCACCTGCAAGCGTTTGGTTGTGGGC  
GTAACCACCTCCGCCATGACGAAGGGTAAGACGGGCATGAATTGGCAAATA  
AAACGCTTATCTTAACGACCATTCTTATCGCTGTCTGCAGGAAAGACGCTGCC  
GGACTTGATTTTGCCCGTGGAAGAGCGCATCGCCCGGCTAAGGGAGTTCCTG  
GTGGACATAGATGATACGCTGCAGTACGAAATTGTGCCCATCGATGATCCCT  
TTGGTCCCACGCAAGTGGATCCTGACCTGGACATGATTGTGGTTCAGTGCAGG  
GACGTTGCGAGGAGGGCAGAAGGTCAACGAGGTACGCTCCGCTAAGCAACT  
GCGCGAGCTGGAGATCTTTGTGATTGACATTGTTGAAAGCAACGTGCATGAT  
GGCATCCACGAGACCAAGGTCAGCTCGAGTAACACACGCATCGATCTGCTGG  
GAACCCGCTGGAGAAGGCCGGAGCCACGACCACAGCTCCCGCCGCGCCCTTA  
CATTATTGGACTCACTGGCGGCATCGCATCTGGCAAGAGCAAGATGGGCGAG  
AGATTGGCCAACATGGGCGCCACGTGATCGACTGCGATAAGGTGGCGCACG  
ATGTTTACGAACCTGGTCAGTTGTGCTACACCCGAATTGTGCAGCACTTCGGA  
CAGGGTATTGTTTCAGACGATGGTTCGATCGATCGGTCCAAGCTGGGACCCTT  
GGTGTGTTGCCGATCCCAAGCAGTTGCAAGCACTCAACGGCATTGTCTGGCCG  
GAACTTATTGCGGAGGTTAACAGGCGGCTGGATGCACTGCGTTCCCAGGCGG  
ACGTGCCGCGTGTGGTGGTCTGGAGGCAGCGGTGCTGCTGCGAGCGGGCTG  
GGAGACCAATTGCCATGAGGTGTGGTCCATGATTGTGCCACCGGATGAGGCT  
GTGCGGCGGATTATTGAGCGCAACAAGTTGAGCGAAGTGGAGGCCCAAAG  
CGACTGGCCAGTCAGGTGCCCAATTCTGAGATCGTGGCCAAGTCGCATGTGA  
TATTCAGTTCGCAATGGGATCACGAATTCACCCAGAAACAGGCGGAGCGTGC  
GTGGAAAATGCTTACCAAGGAACTGGACTCTTACCAGAGCAGCCTTTAACCC  
GATGGATATTTAGATTATCTTGTTGATCCTTATTTTGTATGATTTTTTATGCAT  
TTGTTGTATATTGTTTAGTTGTAAGTCCAAAGTTGAAAAGAAATGCTGGGACG  
TCATTGGGGAAAAACGCTGAAAATTTCAATGGAACCTTAGTGGCTCTCGCCC  
TTCTTGCCAGCCACTCGCTTGAAGTCGTTTCATCTTGGTGGTCATGATGGGGGA  
ACCGATGAAGCCGATATAATCAATCTGCGTCACATCGCCACCTGACTGATTGT  
TCTTCACGAAGATTTGGATGTTTTGCACATTCTGGAACCTTGACGTAGCGCAGA

FIGURE 16 (Continued)

TTCACGGGCACTCCACTCTCCAGCTCCTTCTGAGCCAGGCTGCAAAATGGATT  
GAACAGTGAGAAGAGCTAAGCAGCCATAGAGAAGGCAATAGCTACCTTAGA  
TCCTGCACACTGTTTCATGGACTCGGCCATGTCAAAGTCAATCGTGCGGGGCTG  
GTTAATGAACAGCTTCACATCCTTGGGACCCAGGTGCGAAGGTGCCTTGAAC  
TTCAAAGAGTGGATCTTCACAGCCTGATTAAGGTGATGGACAGGATGAGCT  
GCTCATCGCAATCGGACTGCAGGTAGCCACCGGCGGAGGCCAGGGCGTGCTT  
TAAGTTGTGGTCATCAGCTTCGTTGAGGCACTCGCACTCCTGCTTCGAAATAA  
ATGTATTCAGTTCCATCTGTAAGAAGGATTAGGGATTATTTTGGAACATTTC  
CAAATACTGCACTATATTACCAATCCCTGCCCCTAATCCTCGCCCCCTCCTC  
GCCACCGGATGTACCGATGTGCTCCTGGATCTTGGCCTCGAGCCCATTGACGT  
CCGCACCCTGGACGCGATCGATCTTGGTCCTGTTCTGTAGAAGATGAATGTT  
GGCATGGCCGAAACGCCCTGTCCAGCAGCCGTGTCCTGGCACTTGTCCACAT  
CCACTTTCAGGAAGATGGCCTTTGGGTACTTTGTTGGAAACGTCTCGAAGATG  
GGCGCAATCCGCTTGCAGGGACCACACCACGAAGCTGTGAAGTCCACCACAA  
CCAATTGAATGCCCGCTTGGGCCAACTCCGCCTGGAAGTGGGACTCGTCGTT  
GATCACGCGCACGGACATGGTGATAGGATTAGGTTTCTATTAATTGAGCTTTT  
GTTTCGGCAGCCGAATTGGATTTAAGCAAGTAAATGTTATTATTAACGTTCAA  
TGCAAATTTTTTTTGTAAAGATGACTTGTAATATGCATTTAGTCCAAATTCGT  
GCTAAGAAAAATACCGAATGCGGTATTCCACAAGCGGTACACTGTGATGGT  
ATCGATATTTGAGCTCTTTGACTTCCTATTTTTAGAGGGACCATTATGTGTA  
ATAGAAAAAAACCGAACTTAATATTTAACTTTTATTGAAATATTAGTGGA  
TTACAATATGTAAACTATGAAATATTCTCATTGATATAGCTCAAAGTGTTA  
TTTAAATTCATTCAGTGTTTACGACTAGCAATCTACGCTTTCACGCTCATCTT  
AAGCTTACCGCCCATTTGCCAGGGTTGTCAAGGCGAATGAGCGGTCCCACCA  
TACACGCCACTGGAACCTTCGATACCTGCGCTGCGCCTGGCCACACGTTTATT  
ACCTCGTGGTGTTTCAGTCGGTCGCATTTTCATTAAGTCGCCATTTTAAATTT  
ATTAGAGTCAAGTACAATGGCAGATGTGGAAAAGGAGCCCGAGAAGACCAT  
CGCCGAGGATTTGGTGGTGACCAAGTATAAGTTGGCCGGCGAAATCGTCAAC  
AGTGAGTATTCCTTGGCCGGAAACAGCGAACGCTGGCCGATTCTTGAGTCG  
CTGCTACGTGGCGCTTACACAATGCACCGAATGCCGCTTTCCTTGTGCGCCA  
CGCGTTGGTTAATCTGCCTATTTCTGGACTCTGTCTGCTCGTTTAATTTAGAA  
ACCTCAAGGCGGTAATTGGACTCTGCGTGTTGATGCCTCCGTCCGGGAGA  
TCTGCACCCAGGGCGACAATCAGCTCACCGAGGAGACCGGCAAAGTAAGTG  
GTGGCCACCTGGCGGTTCATTCGCGCCAATTTTCATGTCCAATGATTAAGACTTA  
CACCTTTGAGGGTTTCCCGATGGCGAGCCATGTGCTGTGCGGGCTGGGGATC  
ACCTCGTGGTCGCCAGGCGCACGCGGGGACTCCAATGCTCCACGTGCCCGGC  
TTGTGTGCTCTCCAAAAGGTCCCGAGGATTTACAGATTATGAGATCTGAGGA  
CACACCGCGCACTATCATTGATATATAGTACAACGAACAAGCAATCTAATGC  
TTTTATCGATCTTTCACAAACAGGTATACAAAAAGGAGAAAGACCTGAAGAA  
GGGCATTGCCTTTCCACCTGTCTGTCCGTCAACAACGTGTGTCTGCCACTTCT  
CGCCAGCCAAAAACGATGCTGACTACACGTAAAGGCCGGTGATGTGGTCAA  
AATGTAAGTTGAACCTCCTATTCCACATATACCGCCACTAAATACGTAACATT  
TCTTTTCTACAGCGATCTGGGTGCCACATTGATGGTTTCATTGCCGTGGCCG  
CTCACACAATTGTGGTAGGCGCTGCTGCGGATCAGAAGATCAGTGGTCGCCA  
GGCCGATGTCATCCTCGCCGCCTACTGGGCTGTCCAGGCTGCCTTACGTCTGC  
TCAAGTCCGGCGCCAATGTGAGTCCTCCTTACTTCTAGGTAATCCTCCGTTA

004466-20000000

FIGURE 16 (Continued)

ATCCCTGCAAGAAACGGATTGTCTGCCGCGATTCTCCAGCGACTGAACATCTC  
AACACTTGCAAAGATCAGCTGTGGCAGCTGGTAATTGCCCTGGCCTATTATTC  
AGGACTGGAGGCTTCTTGTCACTTGTCCACAAGGTTATTTCTTCTGCAGGCAA  
CGGATTGACTGCGCTCAAACCTCTGACACAGATCAGCTCAACACCTGCGGATA  
GAAACTGTGTCAATTTCTGTAAGTGAACAAGTTCATTCCATAGAAGTGTTTCGG  
TCTTTAAATTTGTCCACATCTCCAGTTTATAGATATGTCGGAATTGTAATCTGC  
AGGCAACGGATTGTCTGCTGCCTTAACTCGTGGCTCAGCACAGCTCAACGTCT  
GCAGAGATCAACAGTGTGCTGATTTCGTGAACTGAACAAGTTTAGATACTTGAA  
ATGTTTCGGTCTTTAAAGTTGTCCACAATCGCAATGATAATGCCGATCAGTTAT  
TGTTATTTTTCGTTATCTATAGTATACTATGATATTTGATTAAAGATTAGTCAAA  
GGGAATTGGAATGTTTTCTTTATCTCTGCTTTGAACTATTTCCATTTTATTTCA  
TACTTAATATTTATGTTTCAATTCTGTATCCTTACAGAACTACTCCCTCACCGA  
TGCAGTGCAACAAATCAGCGAGTCGTATAAGTGCAAGCCCATTGAGGGCATG  
CTCAGTCACGAGCTGAAGCAGTTCAAAATTGACGGCGAGAAGACGATCATAC  
AGAACCCCAGCGAGGCGCAGCGCAAGGAGCATGAGAAGTGCACCTTCGAAA  
CGTACGAGGTGTATGCCATCGATGTTATCGTCAGTACCGGCGAAGGAGTGGT  
TAGTAATCCATCAATAGACACTACATCTCCACTAATTTGTTTCGATGATTAAAA  
ACACGCGCTTGAGGCTGACTTTGCTGGAATGCGGTGTTTGTTCGAGAGTGA  
CTTGTTTTCGCTCGGCGTTTTTTTATACTAAAATGCGGCACGTGCAGACACCAAG  
TTCCGGCTGGCTGTTGTCCGAAGATTGCAAGATTATGAGATCTGAGAACGCC  
AAATTTAAGCTGGATCCTGGATCATCGCAGCCAGAGCATTATTGCTAACATTA  
TTCGTATTCGTTGCAGGGACGCGAAAAGGACACCAAGGTCTCAATTTACAAG  
AAGTCTGAGGAGAACTACATGCTCAAGATGAAGGCGTCCCGTGCTCTGCTGG  
CAGAGGTGAAAACCAAGTACGGAAACATGCCATTCAACATCCGCAGCTTCGA  
GGAGGAGACCAAGGCCCGCATGGGAGTTGTTGAGTTCGCTCGGCCACAAGAT  
GATTGAGCCCTTCCAAGTGTGTACGAGAAGCCATGTAAGTGTGATGCATAT  
TATTATTAATCCTATTCCCTATTATGCGAGTTGGCAGAACTTAATTCCGGACC  
TGGTACACCTTCGGGTGCTAAGTGCAGGCCAGACATTTTGCCAGAACAAATTC  
CAGGCATTGTGCTCTTCAGCAGTTGCCTCAGTGTGGCCTCTGTCTGAACATGG  
CACTGTCAACAATCGTATCCAATCTATTAACCTGTTTTCTTATACTTATTAAAGT  
TAATTTAGAGACTAACTAGTTTGAGCAACCTTTATAAAGTTCGAATTTTAGC  
CGGAAGTAATAGCAAAGTTAAACAATCCTTTTCCTTATCTTGCATTACAGCCG  
AGATTGTGGCGCAGTTTAAGCACACGGTTCTGCTCATGCCTAACGGCGTCAA  
CTTGGTACACGGCATCCCATTCGAGGCGGAGAACTATGTGAGCGAGTACAGT  
GTTGCGCAGGAGGAGCTCAAGGTAAGCTGCAACAATTCCTTGTATTCACGA  
TGCGTACTCAATGAAATCTCAACTTTTTGCAGACTCTGCTCGCGCAGCCTTTG  
GGTCTGTGAAGGGCAAGGGTAAGGGCAAGAAGGCAACAGCTGGGGCGGGCG  
ACAAAGGTGGAACGGCGCCCGGCGTGGAGACCAAGGCATAGACCAGCCCG  
CTGATGATGATCCGCACCGCCAAGCCATCAACGGAAACACAATGTGAACAAT  
TGCGCTGCCCCAACGCTGCGCTCCACAGATTTTACTATCGAATTCGTTGCGTA  
TTAGAGGACCTTTTGAACAACAGAACAGGACAGAAAGAGAAGACGGCAACAA  
TTTGAGGATACATTTCCCCAGAAATCCTCCATCCATCAACAAGGCGGGCGGT  
CGGTTCGGTCCCGCGCCAACCTTTACCTCTTTATTTCTTTACTATAAGCTGCCTT  
CGTTTATCGGTCTGTTCAACATCATCGCAACGAAAAAGCAAAGCAAGAACTG  
TCATCAAATTGTAACAATTTTAACGCTAAATGATCTTAAAATATAATTCAAGT  
GAAACGTTATTAACGCTGCGTAGTAGGTATTAATAAAATTAACATTTTCTAT

## FIGURE 16 (Continued)

AAAACAGCCGATAAATGCCAAACGATTTTTTCATTTATTTACTTTCCGCTGGCG  
CCCAATTTTAATTCGATTTTCGATACGCTTCTCATTCTAATAAATGCACTTGCG  
AGTTGTGTTTATTTTATACGTTTAATTTAGTTTTGATGTTACATTACATTAT  
ACAATTTGTAATTTAGATTTCTTGCCTTTTGTTATTTTAAATTTTACAGTCTCA  
TCTTTGAACTCTTGTATTACGAAAGTTGCAAGAATAACTTCGTTATGTTAAAC  
GTCACTTAGTGCTGTGCTCACTTGGCCACCCAGTTGTCCATCCCAGATCCAA  
TCCCAACAAGACCAGACCAATTCGATGCCGTATACGGCGACTTTGCCCAACT  
CGCTGACCTCTTCCCTTGCGTCAAACAAAATAAAGAACAACAAAAAACGCAA  
TTGCTGCGGATGAAGTATAGAAAACACGAGCAGCACTTGACAGACGACAAAG  
ATATGTGGCCGGTGATCAAAAGAGGATCTGGGATTTAATGGTCTGCCGTGCG  
TTACATACATGGTTTGGTGTACTTTTTTTTTTTTTTTGTTATGATCGCCGCGACTG  
TTTTCTACTCGCCAGACTAATTATTGACATGCACGTCCATCGGTGCGGAGGCG  
GTCACGTTGCTCGACTTCTCCGGAGAGTCCAGGTAAATCTTCAAGGCACGTTT  
CCGGCGCTGCGCATACCGCGTGGTGGACACGCAGCCACCCGATCCAGTCGT  
GCCTTCTCCCTGGCGTTCATCAGGCGTCGCTCCTCCAGCGTCAGTCTCGCGC  
AGGTACCGTCCTATCTCTGTTGAATTCATTGGTTAGTCTAGGAACTGAACTGC  
CACTTGCTCCACGCTTACTTGTACAGGTAGATGTTTCCTGTCTGTGTGCTATTA  
AGCGGATATTTGTCCAGGGTGGTGGGCACGGAGTACCAACTGCTGGACTCGT  
TAACACTTAACGCTGATATGCTTGTGCAGGGGAAGTTGCTGTTCAACTGCAG  
AGAAGACCAATTAGATCAATATACACAGTAGAACGCAATTTTACGAACCTTC  
ATATAGCTCAGTTTGTCTATTGGGATGTCGCTGATCTGTGATAATGAAAGTCT  
GATTTTCGCTGTCTCTGCAGAAGATACGAGTTCGATTTACTGCTTACAGGGCA  
ATATACAGATTTAACTTACGGTCCAAAGTGATTTCTTGGAACCTTTCCAAACTC  
CAGTTTAGCCGGACACCACCGTCTTACAAATAAAGTCAGAGAATCGTCCTTG  
GGCTGCGGGTCCACTTCTCCTCGCACTCTTGACAAATGAACTGAAATGGTGT  
GATATAAAATCCAAGTTAAGTTTTTTTTCTCATCACAGAGACAGGGGAACCCA  
CCTCCGCTGCGATGCTGGAGCGCATGTGCGAGTAGAGCGTTTCATCGTCGGA  
CAGAATTTTGATGGGAATACGTCCGCCCTTCAGCCAGATGCGACAGTTTTGCA  
CAGACAGTGTGGCGTACTTGGCGTCTATGCGATGCAGCTTGGCCACCAGTTCC  
TTCTTGGCCTGCTCCGCCGTGGTGTGGCATTGTAGACCCACTCGCATAACGCA  
AGGCAGTTTGGACGTCTCGTTATCAATGTCCGCCAGTCGCAGGAAGTGAATC  
TTGGCCTTGAACCTCGTCCGGCTCCAGCGTCTTACCCAACTCCACCGTCAGTGT  
TTCGCCTTCTATGAGATGTACCAACGAGTTGTTCTGGTTGTTTCGACAGATTAT  
TATCGTGTTCGCTGCAGTTTAAAGTGGGCGGCGGGCACTTGGATCAGCTGC  
TCGATGTGTTTCTTAAAGGCCCCCATTCGCATGTGTGTGCCACCAGCAGCTT  
ATAGGCACGCGTGGGCTTACGCAGCTGAGCCTCTTCGTCGGACTGGTGACCA  
GAACTGGAACCGGTGCCAACCACATCCACGCACTCGACCTTTGTGGCGTAGA  
AGAAGTGGTTGGTGTGCTGGTGGGCAGGAGCAGCGGATCCACCACATCGGCCGC  
CGCATAACTGCCGTTTCCATTGCAGTAGGCATGAACGCGCATCATGGCATCGT  
GTGAGGCTGCTTCGTCTCCGGACTAGACAGCTGCGGCGAGTGACTGGTGGA  
ACTGACTTGGCTGTGCGCTCCACCGCGATGTGCCATGTTGTCCGTCTCCACCA  
GAGTTCGGTCTCCGTGCTCAGACTACTGTCCTCCGAGTTGGACTCGTGGCCG  
TGGCTGGGACTGGGTTGTGACATGGGTTCCACTAGATCTCGCTTGTAGCGCCT  
CCAATCGTAGTCGTTGCTGGATGACATGTGACCAGGTGCCACGCCATTCATCA  
TTGCAGCATCCACAACATCAACCCCGCTGGCGCACTGAAAGGACAGCCAAAA  
ATTAACCTTAGTTATAAACCCAACAGCTGTATAACCTACCTCGACTGACTCAA



FIGURE 16 (Continued)

CCACCGGTGAGATTAGGTTTCTACCTTGTGGATACAGCTGCAGTTGCTCAGAG  
AGCTGTTGCTTGCTCAGGGTGTCTACTGGTTCGCAGTCCAACCTCTTTGATGGG  
TGATGACGGCTTGATTGGACTCAGCAACTCTAAAGTCGGCTTAGAAGTGACC  
TTGGCCGTTTTCTTGGCCACAGGACTCTCCGAAGAAGTGTTTTAGAGTTTAT  
CTCTGTAGACAGCTCGGGACATTCTTCAGGGCTAGCCCTGGGAGCTTTTTCCG  
AACC GGCTTGAGATCTTTGCTGCAGTCGTCTTGATCTTGGAAGTCTTTTCG  
GGACTTGATTCCGAAGTATGCTGGTCTTAGCCAAAGAGTCCTCACTCGTCGT  
CTTGGCCTTGCTTGAGAGGAAGATCCCGAAGCTGGCTTCTTTTTCGTTTTCT  
CACCAACTACGCGTTTCTTCTTCTCGCCGGTGGCAGGACTCTTGGCCTTCTCG  
CCGTCCGACTTCATAACCTTTTTGACTACCACTCTTTAATGGGTAAC TCAA  
GCGTTTGGTCACGTCACCATCCCAACTGCCGGAGGGCAGCAGGATCAAGTGA  
TTCTTCAGCTGGGGCTCAAAACCAGCCACTTCGTACATCAGCTGAGATTCCAG  
GGCATTCAAGATTGACCTAAAGTAAAGGGGAATTCAATTAGCGGTTTATTAGA  
ACCTCAAGATGTGCAGATATTTTACCAGATCCTTGTTATCGTGTGGCTGCAG  
CAACAGCTCGAACTTTTCGTACGAGAAGTGCCTGCCAATAAGGTCAATCAGC  
CGTTTCACCGTGAAGTGGGAGCGGACCACTACGTTGATCTTCTTTTGCTCCGA  
GCCGGGTGTCTGGTCAAAGACCGAGACGGTGCAGTCTCGCTCTCCTTGTCGG  
TCATGTCCAGCCCGCGGAATAATCAAGTGATGGTGGAGAAAACCTGCAAAA  
AGATTGTAGGCGAAACGTTGGCTTTACTTATGAATTTTGTCTGGAGTTTTCTTT  
TTATTTTTTTTTTATTTCTTTATTTTAGAATTA AAAAGGTGACACGACACCTTG  
ACGTTTTTCGGCGGGGCCAAGTTCCTGGACATGACGATGCTTCTTGCCCCATAG  
TAAATAAGGAAGAGATGCCAGCCCCAAATTACTGCGAAATCTTCTTGTTTTC  
GACCCCATTCGCGAATAAAGCGGCAGAAACCAAGAAGATTCCGTCCACCTC  
CCGCAGCCGCAGATATTGACGTGCTCCGGGTTTGCTTTTCGCGCCTTATTTGT  
ACGGGCCAGCACCAGTTGCCGTATACATATATATATATATATATAGATAGAT  
ATACACATATAGCACGTACACCCAATCGAGCATCGACTGCCCCCGAAATCG  
ACGTCGTGACTAACGCGCAGGGGAATTTCTGTAAACAACCGGCCATCAGAGTT  
GCCTCCGGAGGATGCTACGGGAATTATTATTTGCCTCCAATGGACTACCAAC  
GTCATCATCATCATCATGACCATAGCTATCACCATCGGGCGTACCGAATGCAT  
AAATTTCAGTGCAAATGTCGCTCCATGTTTCAGCTGGCTTCCTTTGTGGCTCC  
CCGCAAGACTCTGTAACGGAAGTGGTGGCTATTATACGAACGAATATCTGGC  
GCCTTCAATTCGGCAGTGCGCATATTGCAAGTGGACGGTGGACATATCCATA  
TGTACAAATTAATACTTATCGGACATCAGCGTGAACACTGCGAATTATTCTAG  
AAACATTTGTAGAATTCGAAAAGATTTAAGGAAAGCAGATGCTGAATATTAGG  
CGAAAAGCGATTGAACTACTCTATAATATGCAGTCAAAAATATCATCGATTCT  
GCCTGTCAATTAATTGTATCTAAAATTATACTTTTCGAATGTCTATTTTGGCAA  
TAATCTTTAGTGATTTCGTACTGCTCAGCATTAAATTGAGTGTGCAAGCAATT  
GGGGCCGGGGTATTTGCAATGTTTTTCCAATTCTCTGCACCGAAATAACCACA  
AAAAAGACAGCCAGTCAGCCAAGATATTTTGGGTCTCCTCCGAATGGAGGAT  
GCACATCCACGATGTGCGATGTGAATGCGCTGCAATTGGGCGTTCAAACACA  
TGTTGGATGGTCCAAACACAAACCGCATTGCCCGGCAAGGGAGCGAGTGAGA  
TGGGGATCCAAAAATGCTAATACAGTCGGCCAGCACAAAATCAAAATAAG  
AAACCCATGCTGCTAAAAATAAAAACTGGCGGCGGCGACACAACGACACAT  
CGGAGCGGTCCGAAAAAGCACACAGGCGAGTGGAGGAGCAAGATATAAGAC  
AGCTTTGGGAGCGTCTTGAATACGCGTATATCTGGCTATTTGTGAATGCGAAG  
GTTTTTGAGAAATTCAGAGAAGCGCACAGACTGTTTCAATACGTCTATCCTAT

001100 20000000

## FIGURE 16 (Continued)

ACATCAGAATGGTCAGGCACTTTCAACACATTGGCCCCATCCATCCCACTCAA  
TATTTACATGATGACGATGATCTTTTGGTCAATGTTTGTGTTGGTCGGGTATT  
ACAGAAACCGATATCGCGAGTTATCTATGCCATATACACGATCCAATGGGGG  
GACGGCGGGAGGGGCAACAGTCATGCTCGCATATATTTGTGCTATTTTTGAA  
CTATTTTCGGTACTGCGAAATCTATGTGATCTACAAAAACCATGAGATGTCTGA  
GATATGACTGCTGAGTGCCGAAATTGTAGGATTCTCGATTCCCGATCATATA  
ATGCATTCTCGAACAGAAAATCTCCATTACGAAATGCTTTCTATTCTTAGGCG  
TCGCACAACTTTAATTGGAGCTTCCAATGTTGTGTGAATAAGTGTGTATATAT  
CCGTGGTCTATATATGCAACGGATTTTGGTGAGTTTACCCTGTGTGTCGGAA  
CTGAGTGTGCCGAAATCTTTCCGAAC TAGAAGACCGCACCGTCAACGCACGG  
CATAGTTCACGCGTGTACTGGCCGCTTAGGATGCCGATGCCGATTCCGATTGC  
GATCCGAAGATACACCACCCGATCTGGCGCCCGATCTTTGGCGAAGCGAGCT  
ACGTGTAAAGTTCTCGGCGTGATGTACTATAACAATGAGAAACAGTTTACTTA  
TCTGGCTTACACTTCAATAGGAAAACAATACTTTTATATAGCTTCTATAACTT  
CGGGGTGCGATAAGAACATGAATACAGATACACGGATTGCAACAGTACCCA  
AGCCACTTGTTTTAAACAAATAACAGGATAATGGGGAGTAATGTAAGCTATT  
GACTGGGTTACAATCAGGGGTCTGATAACAATCAAACATTGTCCAGTTGCCTT  
TTGCGAATATCAATGACCACTCACGAGTTGCAACTGATAACGATTATCGCCG  
CACAATGCAGTGGGTGGGTATTTCACTGGGGGGAACTTTGGGTCCCTAGAA  
CCCAGACGGATTACTCAATGAATATAGGCGATATGTTTGGGTTTACAGCGAA  
AGTGCTATTAATGTGACCGTATGCTCTCTTCGATGTGCCAGCTCTCTATTTGC  
GGGAATGAATGACTATTTTTATGGGTCTGCCGTCGCTGCTACAATGCTGCATT  
GCTGCAGTGGGACATCCTTTGAACAGGCGCCATGCCAAAGGATATTCTTTGT  
GGAAGGGGGGGGGGGGGGCAAGGGTTAAGGGTCAACATTGCTTTGCGCAATAC  
TTCCAGCGATGGGGCGGTGAACGGTGGGCGGGGCGATCGGTCAAGGCTTCGA  
CTGTGGAACGTGACACGCATATGTGCGCCGGAGTTTGGCCCAAAAAGTGGCC  
CCAATGGTTGTCCTTCGCGCTGGCAATTAGTCCCTAGCAAGGCGCGTCCATAT  
TTTGCAAAAATTGCTGGGGCGCCTTGTTTTCTTCTCTCTGTATGTGTGCATGTG  
TGTGTACCTCCGTCTCACTCACCTCAAGTGTGTGTGTGTGTATGAAAATACTG  
CGGTATACGGCTGCGTTTGTGTGTGAGTGTGGGTTTCGGCTCTACTCTCCCGA  
TGATCCTGCTCCTCCGTCCTAATCCCGGCCTGCTCGGCTGCTCCTGCGTCCT  
GACTGCGCTAGAAAATTCGCTTAAAACGAGCCTCGACGGGTCATTTTTACAATT  
GTTTTTTGTTGTTCCGTTCCGGTGTTTTACCAGACGTGCTCGTTCCGGTGTGAC  
TGCCCCGCCGCTGACTGTAAAATACTAAACGCATTGCAGCTGTGGCAATGCCC  
AAGTCTTGGTCTTACGGTCACACTGGCAAAGTTTAAAAATTTATTTATTTCAA  
CTTTCAGTTACTTTTCGTTGGCTTGAATATTACACTAAGAATTCAATTTGACAC  
TTGCAATTTATACATTGTATATTATAATATATTATATGTATTATATTTTATATC  
ATATAAAGATATTTATATCTATTGATCTTTTGATTATAAGCTCTTTGGTTGAAC  
AATATAAGTGCAACTTTCTCCATCACCTTCCTATCTTTTACAATATGCTTACC  
TCGTCAATACGTTTTTTCTATTTCAAATATTTCAATATTTCAAAGAAATATTTT  
GTTTATTTTTCTGTGTGTTTTTAAGCAATCTGACCCCTGTAGAAGAATCCCTTA  
TAATATTAACAAATGTATCCTCAAAATAGATCGATCTCTATCTTCGCAGACTT  
ACACGAAACATTCCAGAACCGATAGTTTTATGCGATATATGAGATTTAAGGA  
GTACTTTCCGCATTTCCGCATCACAGTCACGCTTTCCTTGGCATTGCAATCA  
AATAAGCGCTAATAATAATCGTAAAAGCATAAGAAGCATATAAAGAAGAGT  
CACCGCCAAAAGCATGCACAAATATATATAAATGGGGAGCGATTAAAAACA



FIGURE 17

dTPR2 5' region, 13015 base pairs

AGACAAAGACAGCGCTGACTTCAGTCGACTTTCGTATTCATTGTTAAATGACA  
TGCAAATGTACGAATGACATGGCATTTCGCCAAAGGGTTTTGAAAGGGGGGCC  
AGATCCAAAGGGCAGGTCTCAGGGAAATGTTTCCAGGCTAATTGTGGGTTTT  
ACGCCCTGTACTTCTCCAAATGATCAAGTACGTCATTTAATGGAAGCCACTGA  
CAATTGGAATCGTAAATTATACAGCACAACTAGATTTGTTTGAGTGCTCTCA  
ATGTAGGCTAATATTAGATTTCTGCGCTGAATTAATAATTATTGTAATACGTAT  
TATAATGCATTTGTACCCAAATTTGACAGACTTAAGCAGTTCTCTAACATAAT  
TGGCATCATTGGCAAAGAGAAATAATATTAAATTGGCAGCATTGCCAGAAAA  
AACTCTTCTCCTAAATTTTGCTTGATTGAATGTTGTAGTTGAGAATGTTGTAA  
AATAGTGTTAGTATTGTAACACACGACATTTTCAAATATTTAAATGAAAATC  
ACATGGTAATTAGCAATTTTGGGTGGCCTTCTTCTCCTCCCAAGCCAAAGCCA  
TATAATTTTCAGCCAGCTACTTGCGATTTCCCCCATGACCAACAACAACAGCCC  
CATATGTGCAGTGCATTAATGCAGATTTCTTGGCAATTGTTTTTGCATACTTTG  
TTTTTCTCACTCACTTCAATTTCAATTGGCGTGCTAATAACTCATTTAGTTC  
GCAACAAAAAACAAAAACGAACAGCGGGCCACAAAAAATGTAGCTACAA  
ACATGGCACACCAACAATGGATTGGATGGCTAACCAAGATCGCCCCACTTC  
CCTTCCCATCAATTGCGAATATATCGCATCTCATGATGCTGAGAGAATACTCG  
TACTCAACTATGCCGACTTTATATGAACACTGTGTGCAGTTTTGTTTTAGGCTT  
TGTAATTATTATAAAAAATAAATTGAACTATTGTTGCCTCATTAGATTGAACA  
GTGAGGCAGCCACAATGTTGCTTTTGTATTTCGGATACACTCAATTAAGCTGA  
ATTTGCAAAATGCAATGGCCCGTATGAACTCACACCTCGAAAATCATAGA  
CTCGAATTATTTTAGAAATTTAATAATAATTATTTTGTTCCTCTTTTTTTTT  
GGTTTGTTTTTTTTTTTTTTTTTTTGTTCCTTGCAACACTTTTCCGCCTC  
TCATTTTGACAGCCCGAGGAGTTCGGTTGGTTCAGTTGATCTCTTGATTGTCA  
GTCAGTCATTTGTGATTAGACATTCGACAGTCGCCGCTATTGTTGGATGGCAT  
AAATTATAGTCTGTCTCAACAACAAAGCGCTGCATATGAAATCCACATAATA  
AATCAATGTGCTGTCGTAATTTGTGTTAAGTTATTTGTAATCAATTTGAATTCT  
CGCCGTACCTCCCCACCCCTCGGTTGGTGAGATTTATGGGAATATTTTATT  
CATTTTGCTATTTTGGTTAAATGGCTTTTGGGGTTTTCCCGAATATAAGTTTA  
AAATTAACGCGCAATAGGCTTAAGATCATGTAATATTATATATTGCCCGTA  
AACAAATGCTTTCTACTTTCATTATCATGAGTGTTTTAAAACTCCACGACTGC  
TCTAAACTTTAATCTTTAAATATTTTTGTACCCTTTGAAGAACTAACCACCTAG  
CAAATCCCTCCTATTATTTCTCAAACCTCTTGCACTTATCGAACTCGCTTCCTT  
TCCCCGCCATCTTCACTCGAACAAATTTAACAACAAATTAAACTGAAATGCA  
GTCAAATCAATCGCTGACTTTTCAATTCGTTTTTCTTCTTTTTTCGGCCCAACA  
TTTTCCACTTGGCCCGAGCGTTTTGCATAGTCCATGGCTTCGATTGGATCGGC  
TCGGATCGGTTGGTAAGTCTTCGGCGGAGTATGGCTTTAGTCCAATTTAGTGG  
AAAGGTGTGCCACACAGCTCGGTCACAACACGTTGCTGTGGCTCATTGGAGT  
TTCGCCTTTGCCTCGCTGGCTTTTGAGCCGTTTGGTCGGTGCCGCTTAAACGC  
CGTTTTAGCCAAGTTAGGTGAAAAATGCCAAGGGAGTGAGGAGTGGAGACC  
GAACTGTCAACTGTGATCAAAATCAATTGTTTGCCATTTGCCAAACCAAATTG  
ACTGAGCCAAGTCAGTGCGAGTCACACAAAAATGCTGACAAAATTATACCAT  
AACCCATGAAATGTCAAGTGTCAATAATTTTTGTAATTATGAGAGCATTGAGCT  
TGAGTACATAAAAAAAAAGTTATATATTTAAAAAAATCATTATTTTAGTTGGC

FIGURE 17 (Continued)

TGCCATTGGAGAAGCCCCCAAAAAAGGCAAACAAATATAATAAAAAATTATT  
GCAACGTAAGTTTTGATTTGAACAAAAGGCGTATACAATTGGATGAGCTCAA  
GAGTGTTTTAGAGTGAAAATGTGAGGATCATTGTTTCGCAACCACTAACAGA  
GGTTCGTCTCTAACATTTTTTCAAAAAAATTACATAACTTTTAAATTTGATTTCA  
GTTTATTTGTAAGTGAGAAGCCTATTTTCTAACCATAAATTCTGCACGTTAAG  
AGTATTTCTTTTCATATCGTATCTACAAAAATCAATCCAACACACCTGTTTCA  
TCTACCGTTAACACCGTTAAGCCCCGCCCATTTTCTTATCGAAAATATAGCC  
CTTTTTACGCTCTATTTATAGCATTACATTCTTTCTTTTTTTTTGCACTTTTT  
AGCTGGCATATCCTTTTCGACTTCGCCCATTCGAGGCTCGCCCAATTTCCGTTT  
CGAGTTTAATTAATTTAATAAACAAATTCTTTTCGCTCTAAAACTCTCAAGT  
GTATCGATACGATGCGTTTCTTTTTTCTTCGTTAAATAAATAATAACCAAA  
AAAAAAAAAAAAACCAAAAAGTAGGAGGAGAAAAGTTATTGCCATAGTTTTTTT  
ATTATACTTGTGTGTTTACCTTTCTGGTGGCTTGATCGATAGGCATCTGCAATT  
AAAAAGAGAAGAAGAAGAGACAAGTGAGGCAAAATTGTAAACGTTTTGTG  
TAAGCTTTAATACGAAAAACAAGTACTGCAACATAACGGAAGGAAACACGG  
CTTAAATTCGGGGCACAAATGCTGAAAGGGAAGTTTTTCATTGACGGGTTCG  
TTCTGACGGACTTGCATTTTGGCGGGCAAGCGGGTGTGAAAATGCACACGCC  
CCGAGAACCCCCCTTTCCACCCCCCTTGACCCCTTATCCAGCCCACTGGCC  
AAAAACAATTTGTAATTATCCACAGAGAGCGCTGCCTTCAGCGGTTTCGCATT  
TCCCCTTTCGCTCGCTCTCCCAACTTGTTTCAATTTAGCGCAAACTTTTTCAA  
CCTAATAATAGGTTTAACCGCATTTTTAACC GTTCCTCATGTTCCGGTCCGGTTC  
GGTTTTCAAACCGGGAATCGTACTTAGACTGGGTCTCCTTATTTCTGTTCTG  
GCTCTCTGTACAATTTTCATTGAGAAAAATGTAAGTAGTTTTTTCATAGCAAC  
GGAATACAATTTAATCCAATAATCCAATAGTTTAATCCAATACAAATGATATT  
ACTACCATTTCTATTTTCGTAAATTTTCGATTTGACTTATTTGGCTGGATTTACT  
TTTCAAAATATATGTTATCAATAAGACACAAACCTTACTTTTCTAGCTATTAA  
CATAGTTTAAAAAAAAAAAAAAAAACTAATAAAAATTACGTGAATCTAAATTTT  
TAAACCCGATATCCAAGAAGATCTCAATTTTTGCCTGTGTACTCAGTTCTCTG  
AACAAAGCGCATGTGCACTTTGGAGCACACTCCATACATGTGGCTCAGCCCT  
TTTCATAATTAAGTAGATGGTTTTCCATCGACTTCATTGTGGTCAGCGGCCA  
GTTCAACCCGTTCTTCACTGCAACCGAGAAGTGTAAACACAAAAACCCAGG  
ACTCTACATTGGCTTAAAAAATAAGAACAGAACAGAACCAAAACCAAAAA  
AAGAAGGGATATTGAAATACAAGGTTGTAATCGTTTTCGACTGTTGATGTCTC  
AATGCATGGGCAGTTCAGTTAGTAAATGTTTTTCAAATCTTTCAGGCAGGAG  
ATGTTTAAATATCCATGAAATATTTGGATCTCCTGGGGATCAATCGGAATATT  
AGCCTTTAATTGTGTTGATCTTTTAAGCCTTTTTGTATCTAATCTAAGCCATTC  
GATCTAATCACAATTTATAAATATCTGCATATTTCTGTATAAGTCTGCATCAT  
TTGACGTAAGTCTTTAAGTCTTTTGGCTTAAGTTGCAACTATAAGGAAGTATT  
TATTTTAGAGACACAAATATTTCAAGTCGCCTTCATTTGAACAAATCGGCGAAA  
ATTGGCTAGCTCGCCAACTTTCTGTAACCAAGGACAATGGTTTTATTTTAAA  
CCATTAAAACTTTAGACCCACTAGCTCCTAGATCCCCCTCAAAGATTTAAA  
AAAAAAAAAACACGATACCCATTTCTACTGAACTTCGTTTTTGCTTGTGTTTT  
TTCCACTCGAACGGAAATGAGCTGACAGCGCACCGCACACGTCGATTGCAGA  
AAACATCGGATAAAACAGGAGGAAAAGTTGTGCAAGGTGGAAGTGTGTTT  
TACCAACTATTGTTAGAGGCGTTCAAAGAATTACGCAGCTTTCGGTTAGTTA  
GCAAGGGGTCACCGGGGAGCGTTACGTTTGCATTGCGTATTTCCGCTAAATGT

004760 20220900

FIGURE 17 (Continued)

CATCGGAAAAGGCAAACGGCGAAATGCGAAACGAAAGTTTTTTGATTGCCCG  
TGTTAATCGATATCGATGCACAACTATTTGCATTGCAACCGTTGCAAGAATA  
TGCAAGAAGTTGGGGGCGGCCGCGGCAGGGGGTGGAAGTTGAGTGCGTAAG  
TTGGCTAAAGCGGAAACAGGAAATGAGAAAATTTTGCAGAGCAAACCCCGA  
ACTGGAAATGCAACTAACTGGGCACATGCACTTTGCGAAATCATTGGATAGC  
GTTAAGAAATTTATTTTAAAATTGTAACAACTTTAATCGTATTCAAAGCA  
ATTAAATCCCAATCCAATTCTTATATAAAATCCTTACAAGATTATTCTATTTA  
CTGTAAATCTAAGCAAAAACCTCCCTTTGCAAAATATTTCGCTGCACAGCACA  
GATCAGTGAAATAATCAAATGAAGTCTTGAAATAACGAAAAACCCCAATTG  
CGTGTGGAAC TGCCCCAATGCTTTTGCTTCGGTTTCGTACCTGGCCGTGGTG  
CAGTCCCTGTAGAGGATGTCGAAGTCCTTGACGACAGCAACTTGACGCGAT  
TGACGCCCGGCTGGATGGCGCCCAAGCCGCGTAATATGCGCACCTGTTCCAC  
ATTGCAGACGAGCGGTACGATGTCAAGGCGCTTTAGTTTGGTATAAACCGTA  
TGCAGACCGCCGACCAAGTGCTTCAGGAAGAGCTCGAAGGCCTGCGGCAGGC  
AGAGCATCGTTTCGTTGCTAATTATGAATGCGGCGACCTTCTGACCCCGGTAC  
TCCACCAGCTTGCACTCATTGGCACTGGGATCCGAGGTGGAGATCGGCGGCG  
GTGAGTTGTACGACCTTGGCGGCACATGGTGATGGGCGGCGGCCATCAGTTC  
CAAGGGGGAGGCATGGTGCATCATCTGCAGGGAGTTGAGGAGCCCCAGCGA  
ATGGGGCGGCAGTCCATGGGGCATTCTGGGCGGTAGGCCCGTCGGCAGTCCG  
TTGCCCGATGGCATTCCATGTGGCGGGGGGCTGAGCTGATGGTGTTGCTGCTG  
CTGCTGTTGCTGCTGCTGTTGTTGCTGCTGTTGCTGCATCTGCTGCATCATGGA  
GTGGTTGAGGGAGCTCACCGGACTGACGGCACTGGGATGGCGACTTGGTGAA  
CTGGCCGGTGAGCAACTGGATCCTCGTCCTGTGTGCGAGCTCCGTCCATTTGG  
ACGATCCTCGTTGGCTCCACGCGAACTGTTGCTACCGTCGCGTCCATTGTGCT  
GCTGCTGTTGCTGCTGGGCTGCTGCGGCTGCCGCTGCTGCGGCTGCCTGCTGC  
TGCTGCTGGTGATGCATCAACATGGCGGTGGTGTTTCATATTGCCCGCGGCTCT  
TTCAATACCACTATTAATATTGTTATTTATTGCGCCCGCTTTATTGTTGTTATT  
TTCACTCTGTTCACTTGTCACAGAATCCATACTTCATCATGGCCGACACTTTTG  
TTTATTTACTTTTTAATCGATTCTGTTAATTTGACGTTTTTTCTATCGTGACAAA  
AATTTGACACAAAGTAAGGGAGAAATAGAAAATAGATGGTGAGAGGAAGAT  
AAATAATTAATGAACTCTTAATTCATTTTTTAATTATTTATTAGGCTTCTATATG  
CAAATTCTAAGTGAGCGTGTCTCGTATATTCCCTATCCGCTTATTATTGGCTTTA  
CATTTTTAATACTTCTGTAAAGTTTTATAACATCAAATTTAAATGCAGACCTTC  
AAAAAATTTACAAACGATTTAGGATTTGTATTAGGCTCAGCTATGCTCCTATT  
TATTAATAATCTATTTTTGAGCCAGTTTAGTTAGTTATATGGTAGCTACAAGTT  
TATATTGCTAAATATTTTTTGTAATTAATATCCTAACAAACATTTTACTTACA  
AAGAAATATAGAGAACTAACAGAAAATAGAAAAGTTTCCTTTCAGACATTTA  
AAGTCCGATTATCTTCTAATACCCCCCATAAATAATCCTTTATCAACAGAACT  
ATTGCTTTGCAAACCTTTGCTTTAATTAAGTTTTGGGAAAAACAAGGCAATGAA  
GCTAATTTGGATCCTTACTGCCAATTTGCATAAATATACCTATTGTCAGCTTT  
ATTTGAATAATTCGATATAGAACATAGATTTACCTTTAAGGAGGTCTAAAAGT  
AATTTATAAACTCAACATCACTGACACAAGACACTCGCGCACTTTGCTTTTTG  
AATTTGCTGTGAAATATATACTCTGAATATTTCAAGTTATTTATCCGATTGC  
CCGCTTGTTAATCGAGTTTGAATAACCGTTTTTCGTACTGGAATTTTGGA  
CCGGAGCTGTGTCCGTTTTTCGAGTACCGTACCGACGGATTGTCACCTCAGAGAT  
TGAGAGATGGCAGCTACTCCGCTGCGACGGCGACGGTGGCGTCGCTGCCTCT

001150-20000000

FIGURE 17 (Continued)

GCTTCTTCGCCTTCGACTGGTTCCTCTTCCCCCTCTCGTTCGGAGAAATCAAC  
GAAACGAATTGCATTCGAATGGGAATCGACTGAGAGCGAGACGGCGCGAGG  
CGACGACTGCGAGTGAGCGAGTGAGCGGGCGCTAACGAGTGCTATTTTTTTA  
GCCCACCCACACACACGTACGTACGTACGTACACACGAAGCGCTACCGTTAT  
GTACTGAGAGAAATGCGCGCGCAAAAGTTTTATTGCATTACCTTCTCTTGCGA  
ATGACAAATTCGTAATGAAAGGCGAGTTTCAATTCGATTCTTTTCGGATTTTC  
GTGGCAGCGACGCCGGCAGCGCGGTCTGGCCGAGACGAGTGTGCTTGTATGTG  
TGTGTCTGTGCCTGTGAGAGCGAGCTGGTGTATCTGTATCTGCGATTGTGCAA  
AACCAGAATACGAATACGAGTACGAATACGAATGTCTGTTGCCCGTCCACGT  
CTCGCATTACACCAATACCAGGCCAAAAAGGGGAGTGGTATGTGCGATTGAT  
CGGTGTGTTTGCATCTGTGTATAATTTCTGTGTGCAACCCCGAAAAATACAATGA  
CAAACGTAATCGCTCTCTCTCTCTTGCATTTCGTTTTATTTATTTTTTCAATTCGT  
TTGTGCGTGTGTGTTCTTCAAATCCTCCCGCTCTCTCTCTTTGGAAAAAAA  
ACGTTTTTCATTTCAATTTCAATTCGTTTCAGTCTGAGCCCTCTCTCTCACACC  
ATCTCGCCATCTCTGTGCGACGCTCAGGTGGGCTGCAACCAATAAACACGAG  
CGAGCGAGAAAGCAGCATATTTGCATAGCCAGTCGTACATGTTTGCCTCTC  
GCTCGCCCCCATGGGCGACGCCTTATATAAAACAAATGACAATTGTTTTGGCAT  
TTTGTGTTGCAAAGTAAATTATAATAAATGCATTGCCAGAGAAGAAAAGTAA  
AAAAAAATAGCTTTACTTCGAGTTTTCGCGAGCTGTCTTTGACAAAAAGCATT  
TAATTTCAATTAAAAGTAAATGACAAACTTTCAACGAATTATACTTTTCGGGG  
CAGTGTGCTATCTCTTTCCGTCCCAAGCTTTGATTTTTTTTTGTCAACCGTTTT  
CCGTTTCCCATTCGTTTCCATTCGAGTCCCGTTTATTTGTATTTCTTTTTGTGT  
GCTGATTTCGGAGGAGAGCAGCACTATGGCAGGGCATCTTTCTTCCACTTACA  
CATATTGCGATAATGGGGTTTTTTTTTCGCCTGAGGGGCGTTTCGTTTTTCGGGT  
CTTATAAATAGCATTGCTTATAAATTCTGGCATCGCACCTTTGCCACCTCTAT  
ATGTTTATGTACAATGTATCTGAGAGCTCGGTCATTTTTCTATTATTTGTCTTC  
GTTTCGCCTTCTGCGATTCTTCTCCATAACGATTGCCATTCCGTGCGCGAACC  
AATCGCATTCCGTTCTCTCCATTTGAGAGTTCCATGTACATATTTCTTTCTATA  
TGGAATGGAATACGTCTTTATGTATTGTGTTTGACATGACGTATGAATTTT  
TCTTGTTTCGTTTCGTTTGGGGCTTTTCTTTTGTGGATTTCCTCACCCACTGTCTT  
TTAGGTGACAGCAACCATTTAATATTAAATTGATTGCAAATGTGGATTTCCAA  
CAGCTTTTAGAAAAATATTTTCGGGCTTTAAAGAAGAATTTAAAACACAATAA  
TTATTGTAATGTAAATATTTTATTTTACATCGGTTTGTTTCATTAAAAAATAG  
TTATAAGATTATATTAGATATGAAATTAATATGTAACGCTACTTTTTTTCTA  
AACTGTGACATTTTAGGCTATTTTTTCTTTTACCATTTCCTTATGTCATATGAA  
TTTCATTTAATTATGACATATACATGAATCGCTGGCTTTAAATTCGAATAAGT  
ACATTAAATTTACCAAAAATGACATGCAGAATTA AAAAGTATTCATTCAAAC  
AAATTTGTTTTCCCCCATAAATGGACAACAAAAAGGTACTGCCTCTATCATC  
CAAGTGTCAAAATATGTCATAGCAACCAACTATCGTCAGTAAGAAATGAGTT  
CTACAACATGCAACTTTTTCATGGTGTGCAACTGTGGGCGGGAAGTTTGATT  
TTTCGCAACAAACAGCTCGCTTTGAACTCTGGTTTTTCTCTTTAATAAATGCA  
ACTGATCTAACTATTAAGTAAAATTGTATTTTTTATTAACCACAAGCAAGCGC  
AAAGATGAGTTTATATTCTAAAAAAAAGGAGGGTGATTAATTTCTATTAGTTT  
GGATTACAAATTTGGACTAGGAGTCAATTTGAAAGTCGTTATATCAATAATA  
CTTCTGGACTTTGAAGCGACAGTTACTGTTCCATAACTTCGGATTATCAGCTT  
TGCCTTCACCACATATATAGAGTATTCTCTGGATGTGTGCGAGATTGTATTTTT

FIGURE 17 (Continued)

AAACGACGACTGGATGGCAAAAGTTCAGTGCGCTCGCAGCTATTATGTGGAT  
TATCTGCCTCTTGCTGGTGGCCCTTGTGGCGGCCAGTTCCAATACAAGACTTC  
TAAATGGCATCCTAAGTCATGTGGACAAGGAAGCCAATCCCTGTGAGAACTA  
CTACAACCACGCCTGCGGCCAGTACAACATGCGTCACATCGACGACACCTTC  
TTCGACATTATACAAATGCTGGATCACCAGGTAAACCAGAACTTGGTGAAC  
TAATGGACGAGCTGGAAATGAGTTCTCAATTGCCGGACTTTAATGTATCTAGT  
GTAGATGGCAAGGTCCTTCGTTACTACCTTAGTTGTCGTGGAGCGCCGCGGA  
ATATGGATAGTTTAAGCCAGTATCTGAAAGTGATTTCCCCCGGCGAAGGACT  
CACATGGCCTCAATTTCATTCCGGACGGTAGTTCTTGGCCCCAGGAGAATTTCA  
AATGGCTCAAGGCACTGGCTCATCTGCATCGCTACGGTCTAACTAACGTGTTT  
TTTAACCTTGAAGTCGTGTCAAACCCACGAAATGCCAGCGAGTACATGGTAG  
AATTAAATACACCCACTTTTGGAGAAGAATCTCAACTGCCGAACAGTTTTATT  
GAAATTCTATCCGTTCTCTATATCATAAAGGTTCCCTCCAGTGAAATCATTAC  
TCTGGCGCGAAAAATGCGAACGCTTGAATTGTTGCTTAAACGATGATCAAT  
CCGATCGACACACTGAATAATAGATACATTAGTATCCGCGATTTTCAGATGG  
AAACCGGTCACAACCTGGCAGCGTTTCTTTGAGATTTAATAGGCTCCAGCGCA  
GCCCCAGAACTCCAAGTGTTGGTGCACAATTTTAGGTACTTTACCGCCCTTAA  
GGAATAATGGACAAACAGGATGCTCGGCTGGTGGCCAGCTACATAATGACC  
CGATTTGCAATATTTCTATTGGATGAAACCATGGGTGGCAGAGAATCCACGG  
AGTGTGTGTACAGGTGCGCCGCAACATGAATTTGGCTGCAAACATGCTCTA  
TAAGGAACGATTTTTTCGAAGACTCCACTTTTCAGTGCCAATATCCTGGAAATTA  
AGGACATTTTCGAGAACTACGCCATCAGTTTCTGCTGCAAGTCGATCAAAA  
TCATCTAGAGTTGACTGCTTTGCAGATGAAATTTTTTCGTTGAAAGGCAGAGG  
CAATTGAGATCAACGTTGTGAATCTTCCAAAAACCGATGATCTTCGCCATTTT  
ATCGGCCAGTACTACCAAGACTTGCAGTTTCCCACTGGCGAGCTGGATTACC  
ATCAGGAGCACCTCAAGGTGCTGCAGTTTCGCACCCAAAAGATGTTGGCCCA  
ATCCAGCAAAGGGCACTCAGAGGAGCAGAATATTTTGACTTACAGGAGCCAA  
GCGGCGCCATTGCCTCCACCTCGTACTATGTGATGCGCCCCAATGTGATTATT  
GTCCCCCTTGGGCTACTGCAAGAGCCATTCTTTCAGCTGGAAAGCGAAGATG  
TCTTCAAATACAGCCTGATGGGATATATTATGGCACATCACTTGATAAGCGCC  
TTTGCCACCGAGGGCATTACAATTGGCAGCGATGGAAACGATCAATCATTTA  
GATCGCATCGTTTTGAAGAAGCAGTCAGTTGCTTGTACGCAATTCAGAGAA  
CATCGATGAAAGCATGGGCGATATTGCTGGTTTAGAACTGGCCTATTTTACTT  
ATGCTAAGATGGCCAAGAATCGAAACCGTTTGGATTTACCCATTTGCCACC  
GGAGCAGATATTCTTCTAAATGTTGGCCAGTTCTTCTGCGGCAATAGCGATA  
TGTTGGTTCAGTACAAGGAAGATCAAGTGCGTTTACAGCGAGCTATTGAAGG  
GTTTGAGCCATTTGACAAGGCTTTTGGGTGCTACCGCAATAAGCCTAAGCAC  
GAGAAGTGTGCTTTATAGTGAATACCTTGACATATGCTTAGAAATACATATT  
TTTTGATAACAATAATACAAGACAATCGTGTTAAATTATAAAAGTGTTACAAT  
CACATCCATTCTGTTCTTTTAAAATTAGTTTTTAAACTAACAATAGTCAATAGG  
CTAAGATAGTTAAATGATCATCATTCGAATAAACAACGTTCAAGATTGACTCT  
TCAATGTCATGCACCTGCAAGATTACCATTTATTATAAATTAAAAAACACAC  
AAAGTTATACGTGTGTTACTTACATGCATTACATTCGGGCCTGGCCATCCACT  
TAATATACTGAGATGTAGCGGTCTTTGATTTGCGGGATCTCTTATGGATTTTA  
GAACATTGTAACTTTGCTGACAAAGTAAATTCAACTTTTAACGACTTGTGGT  
GTGTGCGGCCCGATGAAATGTCTTAAAATACAAATTAAATACAATTCAAATA

## FIGURE 17 (Continued)

TAATTCAGACGTCAAAAGGTTTAAAGTTAAAATATATTTTACCTTTTAGTGTT  
ATTTATACGTATGAGCCTTGAAAACACAGTTGAATATCAAACGGATTTTGTGTT  
ACCAACAGATTCCAACAGATTCTCCAACCTTTCGTTTTTTGATTGCCTATTCACT  
CGAAGATCTATTTCCAGTACTATGATCCTCCATAGTAGAGTCAGCTCAGGATC  
TTGTGATAATCCGCAAGCAATTCGACAAAGAATTCGTCGGCCAGAACAAAAT  
TTATTAAATCATTGTAGTCATTCTCAGGATCTCTCTTAACTGGCAATCCGTAA  
TAACGTATTTTATTATCTCCAAAATACAGTCGGAATTCAGATTAAATTTGCCG  
TTTCCGTCCTTTTTTATAAATATACATACAAATATACTAAGCAATAGACTGAA  
ATGAATTCTAGAATTTGAGGAAACTAATTATGTACCTTTATGAATACTTTTCC  
TACTTGTACTAATCAAACCTAATTTTTAACAGATTTTTCATGCCGAATGATTA  
CAATCTTATTTGGATGATTTGATAGAGCTTAGGAATAATGGTTTTAATTTTGG  
ATTAAAGAGTTGCGATTAAGAAACGAAGATATTATCTAGTTTTTTGAAGAACA  
CAGGGTACTTTAAATTTTCGCACGCGGAACGTCAAAACAAGAAGAAGTTTTCA  
TCAACACTGAATTTCCGCTTGGTAATCAGCTGATAAGCGTGCTCACGATAGCC  
GAGTTCACATCCAACAGATGTTTCCCTTAGCAGGGTTTCAGACCCAAATGATG  
ATTTATCTTATTTTGATTAAGCTCCAACACGCATTGCTTTGCATAATTCAGGTA  
TTATTAGGCTGCTTAATATACAATCCACTTATATTGTTGTGTCCATGAGGAAC  
ATCGACACGTGAGGATAAAAAATATTTATTTATCGATATATTTTTACTCTTGAG  
CCTTTTGCACACCCCTAGTTGTGTTCCA

## FIGURE 18

dMLF 5' region, 8374 base pairs

GCCTGATTGTTTTCCACTTTGCAGCAGAGGAGCCGGGAAGGAGCGGTAGAGG  
CGCACCCAGTGTATCCGGCAAAGGCAAGTCACCCCAAGGTGCGTTCCATGCCC  
AGCTCTCCGCTGCCTCAGCGATCCGCTACGCCGACGCGGCTGATGAGCCAAC  
GTGTCCGTGAGGCGGCCGAGCGTCTTGCCCAACAGCACACGGTGGCCAGTGC  
TCAGCGGCATTTGGGCAATGGGAGAGGCACTGGCACTGGCAATGGAAATGGC  
AATAGCAATAGTAATGGCAATGGTAATGGGAACACCGCGGAGACGAATCGC  
GAATCACGCGCGCGACGTCTCATCAACCGATTCAATAGCGAAACGCAGCATA  
TCACGTCCTAGTTTAAGTCGGTTAAATGCCGACGAGCATAACTTTATTACAGA  
TAAAGCAGATATAGCATTGTTTAAAGTAAAAAATATATATATATACCCAGAG  
AACTTTACGAAACACTCGAATATGAATGCGACTGCGGATCAGCATCCCACC  
CACCCACACACACACGTCTACCCACTCACAGTAGGATATATGTATGTATGTCT  
GCATTCAAGCGGATGCACTCCCTCCGTTTCAGAGGGAAGTGTACTTAGGCTAG  
AGGAAGCTAAGTGTTTAAATTATTGTATCGATTTATATACATATTTACCATAC  
TAATTAAAGTTAATGTAACGAAAACGCAGGATCAGTAATCTTATTTAGTTCA  
ATGGTAATCAATGTGCGATTAGCGGATGATCGCGCTCCTTGAGTCGCACCCA  
CAGTCCGCCGGAGGCTCTCAGCGTAATCCGGAAGGTGGCCGCAATGGTTGTC  
TTTCCGGTTACAGGAAGCAGCTGGTAGCTACGCAGCAGGCGCGACACTATGG  
TCTTGATCTCCATGATGGCGAATCGATTGCCAATGCAATATCTCGGTCCAGCG  
CTGAAGGGTAAAAAGGCGTAGGGATGACGGTTCTCGGAGTTCTCGGGCGAAA  
ATCGCTCCGGCTGGAACCTTTTCCGGATCGGGATAAATGTGGGCAAGACGATG  
GGTGGCATAGGGGCAAATGAAAACGTTGCTGCCGGCGGGCAATGTGTGCTTT  
GCCAGGCGAACCTCTTCGCCCAGTTTACGAGCAATAAGCGGGACACTGGGAT  
ACAGACGCAGTGCCTCCTTGATGCACATCTCCATGTAACGCATCTCGTGCAGA  
TCCGTCATCGTGGGAGCTCTATTACTGTCCTCGAATATGGTCGCCAGCTCCAG  
GACACAGCGATCCTGGCACTCGGGATTCTGTGTCAGCAGAAAAGAGAGTGAAG  
GCCACGGCGGCACCCACCGAATCCTGGCCAGCCAGCATAAAGGTACAGGCCT  
CGTTGACGATATCCTCCTCGGTGAAGTCCCGATTGCTCTCGGAGATCTCGATC  
ATGTGGTTCGAGCAGACACTTTCGCTCGCTATTGCCATTATTATTGTTCTGGAT  
TTGGCGACGTCTCTGGATCATTTTTCGCTGTGAAGTCATTGAGGCGCTTCTTCT  
GGTTAAGCTCATCGTTGGCCATCTTGGTCCAGTGGTAGATCCCGTCCAGCAGC  
AGCCAGGGTTGCGTAAACCGCGCGGGCATCATGATCTTGCCCTGGCGGAACG  
GCGAGTCCTCCATCATGGCCACATCCTGACCTCTTTTCTTGATCGGCACACCC  
AAAACGGCCTCTGCAAGTCGTTTCAGGGATTAAGTGAGAAATTATAGCTTGCT  
AATCCCCTAGAGACTCACCATTTAGTATGTCCAGTACACAGTTGTTACGTAC  
TTGGCAATATTTATCTCCGTTCCACGGCTTCGGCATCCAGATTCTCGTACAA  
CGATTGCGAGGCATCCACAAAGGTGTCGATGAACTTCTCCAGCAGATTGTGA  
TGAAACGCTGGCTGGATGAGCCGTCGATGATTGCTCCACTTGGAACCACTGC  
TGGTTATCAGCCCATCACCCAGGAAATTGTGCATCAGTCGGTAGAAGAAGAC  
CTTGTTGGTGTGCTTCTTCGAGGAGAGTATCACCTGCAGATCCTCCGGCTCCA  
GGACAGCAAAGAAGGGAAAGAGCAGCACCCAGATCCGCACCAGAGATCCAT  
ATAGATCGAAGGCCTTGCCGGCACATCTGCGCATCACTGTGAAATGGGATTC  
AATTAACTTAAAAGGTATCTTTCACGAAAAGGTTTCTTCAAGGATCTTACAG





